

Construction of Slepian-Wolf Source Code and Broadcast Channel Code Based on Hash Property

Jun Muramatsu and Shigeki Miyake

Abstract

The aim of this paper is to prove theorems for the Slepian-Wolf source coding and the broadcast channel coding (independent messages and no common message) based on the notion of a stronger version of the hash property for an ensemble of functions. Since an ensemble of sparse matrices (with logarithmic column degree) has a strong hash property, codes using sparse matrices can realize the achievable rate region. Furthermore, extensions to the multiple source coding and multiple output broadcast channel coding are investigated.

Index Terms

Shannon theory, hash property, linear codes, LDPC codes, sparse matrix, maximum-likelihood decoding, minimum-divergence encoding, Slepian-Wolf source coding, broadcast channel coding

I. INTRODUCTION

The aim of this paper is to prove theorems for the Slepian-Wolf source coding (Fig. 1) introduced in [26] and the broadcast channel coding (Fig. 2) introduced in [3]. The proof of theorems is based on a stronger version of the hash property for an ensemble of functions introduced in [17][18]. This notion provides a sufficient condition for the achievability of coding theorems. Since an ensemble of sparse matrices also has a strong hash property, we can construct codes by using sparse matrices where the rate pair of the code is close to the boundary of the achievable rate region. When constructing codes, we employ minimum-divergence encoding and maximum-likelihood decoding.

The achievable rate region for Slepian-Wolf source coding is derived in [26]. The technique of random bin coding is developed in [4] to prove the Slepian-Wolf source coding theorem. The achievability of the code using a pair of matrices is studied in [5]. The constructions of encoders using sparse matrices is studied in [20][27] and the achievability is proved in [16] by using maximum-likelihood (ML) decoding. In this paper, we construct codes based on the strong hash property, which unifies the results of [4][5][16].

To construct a broadcast channel code, we assume that independent messages are decoded by their respective receivers with small error probability. It should be noted that we assume neither “degraded” nor “less noisy”

J. Muramatsu is with NTT Communication Science Laboratories, NTT Corporation, 2-4, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0237, Japan (E-mail: muramatsu.jun@lab.ntt.co.jp). S. Miyake is with NTT Network Innovation Laboratories, NTT Corporation, Hikarinooka 1-1, Yokosuka-shi, Kanagawa 239-0847, Japan (E-mail: miyake.shigeki@lab.ntt.co.jp).

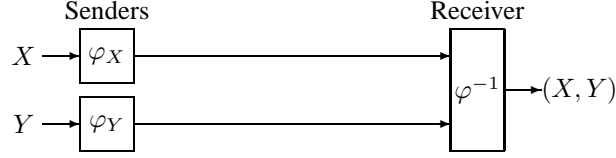


Fig. 1. Slepian-Wolf Source Coding

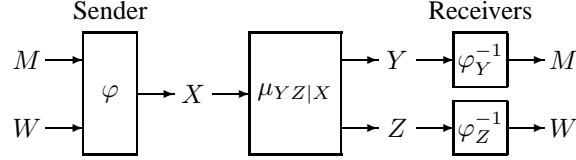


Fig. 2. Broadcast Channel Coding

conditions on this channel. The capacity region for this channel is known only for some classes. An inner region on the two receiver general broadcast channel without a common message is derived in [15] and a simpler way of constructing codes is presented in [7]. The cardinality bound is investigated in [10][11]. Outer regions are derived in [13][21][22]. Applications of sparse matrices (LDPC codes) to broadcast channels are investigated in [2][23][24][25]. In this paper, we construct codes based on a strong hash property and we show that the rate pair of the constructed code is close to the inner bound derived in [7][15].

The proof of all theorems is based on the notion of a hash property, which is a stronger version of that introduced in [17]. It is the extension of the ensemble of the random bin coding [4], the ensembles of linear matrices [5], the universal class of hash functions [6], and the ensemble of sparse matrices [16]. Two lemmas called ‘collision-resistance property’¹ (if the number of bins is greater than the number of items then there is an assignment such that every bin contains at most one item) and ‘saturation property’¹ (if the number of items is greater than the number of bins then there is an assignment such that every bin contains at least one item), which are proved in [17] and reviewed in Section III, are extended from a single domain to multiple domains. The extended collision-resistance property is used to analyze the decoding error of the Slepian-Wolf source coding and the extended saturation property is used to analyze the encoding error of the broadcast channel coding. It should be noted that the linearity of functions is not necessary for a strong hash property but it is expected that the space and time complexity of the code can be reduced compared with conventional constructions by using sparse matrices. This is a potential advantage of our approach.

II. DEFINITIONS AND NOTATIONS

Throughout this paper, we use the following definitions and notations.

The cardinality of a set \mathcal{U} is denoted by $|\mathcal{U}|$, \mathcal{U}^c denotes the complement of \mathcal{U} , and $\mathcal{U} \setminus \mathcal{V} \equiv \mathcal{U} \cap \mathcal{V}^c$ denotes the set difference.

¹In [17], they were called ‘collision-resistant property’ and ‘saturating property,’ respectively. We changed these terms following the suggestion of Prof. T.S. Han.

Column vectors and sequences are denoted in boldface. Let $A\mathbf{u}$ denote a value taken by a function $A : \mathcal{U}^n \rightarrow \text{Im}A$ at $\mathbf{u} \equiv (u_1, \dots, u_n) \in \mathcal{U}^n$, where \mathcal{U}^n and $\text{Im}A \equiv \{A\mathbf{u} : \mathbf{u} \in \mathcal{U}^n\}$ are the domain and the image of the function, respectively. It should be noted that A may be nonlinear. When A is a linear function expressed by an $l \times n$ matrix, we assume that $\mathcal{U} \equiv \text{GF}(q)$ is a finite field. For a set \mathcal{A} of functions, let $\text{Im}\mathcal{A}$ be defined as

$$\text{Im}\mathcal{A} \equiv \bigcup_{A \in \mathcal{A}} \text{Im}A.$$

We define a set $\mathcal{C}_A(\mathbf{a})$ as

$$\mathcal{C}_A(\mathbf{a}) \equiv \{\mathbf{u} : A\mathbf{u} = \mathbf{a}\}.$$

In the context of linear codes, $\mathcal{C}_A(\mathbf{a})$ is called a coset determined by \mathbf{a} . The random variables of a function A and a vector $\mathbf{a} \in \text{Im}\mathcal{A}$ are denoted by sans serif letters A and \mathbf{a} , respectively. On the other hand, the random variables of a n -dimensional vector \mathbf{u} is denoted by bold Roman letter \mathbf{U} .

Let p and p' be probability distributions and let q and q' be conditional probability distributions. Then entropy $H(p)$, conditional entropy $H(q|p)$, divergence $D(p||p')$, and conditional divergence $D(q||q'|p)$ are defined as

$$\begin{aligned} H(p) &\equiv \sum_u p(u) \log \frac{1}{p(u)} \\ H(q|p) &\equiv \sum_{u,v} q(u|v)p(v) \log \frac{1}{q(u|v)} \\ D(p || p') &\equiv \sum_u p(u) \log \frac{p(u)}{p'(u)} \\ D(q || q'|p) &\equiv \sum_v p(v) \sum_u q(u|v) \log \frac{q(u|v)}{q'(u|v)}, \end{aligned}$$

where we assume that the base of the logarithm is 2.

Let μ_{UV} be the joint probability distribution of random variables U and V . Let μ_U and μ_V be the respective marginal distributions and $\mu_{U|V}$ be the conditional probability distribution. Then the entropy $H(U)$, the conditional entropy $H(U|V)$, and the mutual information $I(U; V)$ of random variables are defined as

$$\begin{aligned} H(U) &\equiv H(\mu_U) \\ H(U|V) &\equiv H(\mu_{U|V}|\mu_V) \\ I(U; V) &\equiv H(U) - H(U|V). \end{aligned}$$

A set of typical sequences $\mathcal{T}_{U,\gamma}$ and a set of conditionally typical sequences $\mathcal{T}_{U|V,\gamma}(\mathbf{v})$ are defined as

$$\begin{aligned} \mathcal{T}_{U,\gamma} &\equiv \{\mathbf{u} : D(\nu_{\mathbf{u}}||\mu_U) < \gamma\} \\ \mathcal{T}_{U|V,\gamma}(\mathbf{v}) &\equiv \{\mathbf{u} : D(\nu_{\mathbf{u}|\mathbf{v}}||\mu_{U|V}|\nu_{\mathbf{v}}) < \gamma\}, \end{aligned}$$

respectively, where $\nu_{\mathbf{u}}$ and $\nu_{\mathbf{u}|\mathbf{v}}$ are defined as

$$\begin{aligned} \nu_{\mathbf{u}}(u) &\equiv \frac{|\{1 \leq i \leq n : u_i = u\}|}{n} \\ \nu_{\mathbf{u}|\mathbf{v}}(u|v) &\equiv \frac{\nu_{\mathbf{u}\mathbf{v}}(u,v)}{\nu_{\mathbf{v}}(v)}. \end{aligned}$$

For $\mathcal{T} \subset \mathcal{U} \times \mathcal{V}$ and $\mathbf{v} \in \mathcal{V}$, $\mathcal{T}_{\mathcal{U}}$ and $\mathcal{T}_{\mathcal{U}|\mathcal{V}}(\mathbf{u})$ are defined as

$$\mathcal{T}_{\mathcal{U}} \equiv \{\mathbf{u} : (\mathbf{u}, \mathbf{v}) \in \mathcal{T} \text{ for some } \mathbf{v} \in \mathcal{V}^n\}$$

$$\mathcal{T}_{\mathcal{U}|\mathcal{V}}(\mathbf{v}) \equiv \{\mathbf{u} : (\mathbf{u}, \mathbf{v}) \in \mathcal{T}\}.$$

We define $\chi(\cdot)$ as

$$\chi(a = b) \equiv \begin{cases} 1, & \text{if } a = b \\ 0, & \text{if } a \neq b \end{cases}$$

$$\chi(a \neq b) \equiv \begin{cases} 1, & \text{if } a \neq b \\ 0, & \text{if } a = b. \end{cases}$$

Finally, for $\gamma, \gamma' > 0$, we define

$$\lambda_{\mathcal{U}} \equiv \frac{|\mathcal{U}| \log(n+1)}{n} \quad (1)$$

$$\zeta_{\mathcal{U}}(\gamma) \equiv \gamma - \sqrt{2\gamma} \log \frac{\sqrt{2\gamma}}{|\mathcal{U}|} \quad (2)$$

$$\zeta_{\mathcal{U}|\mathcal{V}}(\gamma'|\gamma) \equiv \gamma' - \sqrt{2\gamma'} \log \frac{\sqrt{2\gamma'}}{|\mathcal{U}||\mathcal{V}|} + \sqrt{2\gamma} \log |\mathcal{U}| \quad (3)$$

$$\eta_{\mathcal{U}}(\gamma) \equiv -\sqrt{2\gamma} \log \frac{\sqrt{2\gamma}}{|\mathcal{U}|} + \frac{|\mathcal{U}| \log(n+1)}{n} \quad (4)$$

$$\eta_{\mathcal{U}|\mathcal{V}}(\gamma'|\gamma) \equiv -\sqrt{2\gamma'} \log \frac{\sqrt{2\gamma'}}{|\mathcal{U}||\mathcal{V}|} + \sqrt{2\gamma} \log |\mathcal{U}| + \frac{|\mathcal{U}||\mathcal{V}| \log(n+1)}{n}, \quad (5)$$

where the product set $\mathcal{U} \times \mathcal{V}$ is denoted by \mathcal{UV} when it appears in the subscript of these functions. These definitions will be used in the proof of theorems.

III. STRONG (α, β) -HASH PROPERTY

A. Formal Definition and Basic Properties

In the following, we introduce the strong hash property for an ensemble of functions. It requires a stronger condition than that introduced in [17]. It should be noted that the linearity of functions is not assumed in this section.

Definition 1: Let $\mathcal{A} \equiv \{\mathcal{A}_n\}_{n=1}^{\infty}$ be a sequence of sets such that \mathcal{A}_n is a set of functions $A_n : \mathcal{U}^n \rightarrow \text{Im}\mathcal{A}_n$. For a probability distribution $p_{\mathcal{A}_n}$ on \mathcal{A}_n corresponding to a random variable $A_n \in \mathcal{A}_n$, we call a sequence $(\mathcal{A}, p_{\mathcal{A}}) \equiv \{(\mathcal{A}_n, p_{\mathcal{A}_n})\}_{n=1}^{\infty}$ an *ensemble*. Then, $(\mathcal{A}, p_{\mathcal{A}})$ has a *strong $(\alpha_{\mathcal{A}}, \beta_{\mathcal{A}})$ -hash property* if there are two sequences $\alpha_{\mathcal{A}} \equiv \{\alpha_{\mathcal{A}}(n)\}_{n=1}^{\infty}$ and $\beta_{\mathcal{A}} \equiv \{\beta_{\mathcal{A}}(n)\}_{n=1}^{\infty}$ depending only on $\{p_{\mathcal{A}_n}\}_{n=0}^{\infty}$ such that

$$\lim_{n \rightarrow \infty} \alpha_{\mathcal{A}}(n) = 1 \quad (\text{H1})$$

$$\lim_{n \rightarrow \infty} \beta_{\mathcal{A}}(n) = 0 \quad (\text{H2})$$

and

$$\sum_{\substack{\mathbf{u}' \in \mathcal{U}^n \setminus \{\mathbf{u}\} \\ p_{\mathcal{A}_n}(\{A : A\mathbf{u} = A\mathbf{u}'\}) > \frac{\alpha_{\mathcal{A}}(n)}{|\text{Im}\mathcal{A}_n|}}} p_{\mathcal{A}_n}(\{A : A\mathbf{u} = A\mathbf{u}'\}) \leq \beta_{\mathcal{A}}(n) \quad (\text{H3})$$

for any $\mathbf{u} \in \mathcal{U}^n$. Throughout this paper, we omit the dependence on n of \mathcal{A} , A , \mathcal{A} , $\alpha_{\mathcal{A}}$, and $\beta_{\mathcal{A}}$.

Let us remark on the conditions (H1)–(H3). These conditions require that the sum of the all collision probability which is far greater than $1/|\text{Im}\mathcal{A}|$ vanishes as the block length goes to infinity.

Remark 1: The condition

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \frac{|\overline{\mathcal{U}}_n|}{|\text{Im}\mathcal{A}_n|} = 0$$

is required for an ensemble in [17, Def. 1]. We omit this condition because it is unnecessary for the results presented in this paper.

It should be noted that (\mathcal{A}, p_A) has a strong $(1, 0)$ -hash property when \mathcal{A} is a universal class of hash functions [6] and p_A is the uniform distribution on \mathcal{A} . The random bin coding [4] and the set of all linear functions [5] are examples of a universal class of hash functions. The strong hash property of an ensemble of sparse matrices is discussed in Section III-B.

From the following lemma, we have the fact that the above definition of hash property is stronger than that introduced in [17]. It is proved in Appendix B.

Lemma 1: If (\mathcal{A}, p_A) has a strong (α_A, β_A) -hash property, then

$$\sum_{\substack{\mathbf{u} \in \mathcal{T} \\ \mathbf{u}' \in \mathcal{T}'}} p_A(\{A : A\mathbf{u} = A\mathbf{u}'\}) \leq |\mathcal{T} \cap \mathcal{T}'| + \frac{|\mathcal{T}||\mathcal{T}'|^{\alpha_A}}{|\text{Im}\mathcal{A}|} + \min\{|\mathcal{T}|, |\mathcal{T}'|\}\beta_A \quad (6)$$

for any $\mathcal{T}, \mathcal{T}' \subset \mathcal{U}^n$, that is, it has a (α_A, β_A) -hash property introduced in [17, Definition 1].

In the following, we review two lemmas of the hash property. It should be noted that Lemmas 2 and 3 are related to the collision-resistance property and the saturation property, respectively. These relations are explained in [17, Section III]. Let \mathcal{A} be a set of functions $A : \mathcal{U}^n \rightarrow \text{Im}\mathcal{A}$, p_A be the uniform distribution on $\text{Im}\mathcal{A}$, where random variables A and \mathbf{a} be mutually independent, that is,

$$p_A(\mathbf{a}) \equiv \begin{cases} \frac{1}{|\text{Im}\mathcal{A}|}, & \text{if } \mathbf{a} \in \text{Im}\mathcal{A} \\ 0, & \text{if } \mathbf{a} \notin \text{Im}\mathcal{A} \end{cases}$$

$$p_{A\mathbf{a}}(A, \mathbf{a}) = p_A(A)p_A(\mathbf{a})$$

for any A and \mathbf{a} .

Lemma 2 ([17, Lemma 1]): If (\mathcal{A}, p_A) satisfies (6), then

$$p_A(\{A : [\mathcal{G} \setminus \{\mathbf{u}\}] \cap \mathcal{C}_A(A\mathbf{u}) \neq \emptyset\}) \leq \frac{|\mathcal{G}|^{\alpha_A}}{|\text{Im}\mathcal{A}|} + \beta_A.$$

for all $\mathcal{G} \subset \mathcal{U}^n$ and $\mathbf{u} \in \mathcal{U}^n$.

Lemma 3 ([17, Lemma 2]): If (\mathcal{A}, p_A) satisfies (6) then

$$p_{A\mathbf{a}}(\{(A, \mathbf{a}) : \mathcal{T} \cap \mathcal{C}_A(\mathbf{a}) = \emptyset\}) \leq \alpha_A - 1 + \frac{|\text{Im}\mathcal{A}|[\beta_A + 1]}{|\mathcal{T}|}$$

for all $\mathcal{T} \subset \mathcal{U}^n$.

In the following, we consider the combination of two ensembles, where functions have the same domain. It should be noted that the assumption of a strong hash property makes it unnecessary to assume the linearity of a function for the hash property of the concatenated ensemble while the linearity of a function is assumed in [17][18][19]. The proof is given in Appendix C.

Lemma 4: Assume that an ensemble (\mathcal{A}, p_A) of functions $A : \mathcal{U}^n \rightarrow \text{Im}\mathcal{A}$ satisfies a strong (α_A, β_A) -hash property and an ensemble $(\mathcal{A}', p_{A'})$ of functions $A' : \mathcal{U}^n \rightarrow \text{Im}\mathcal{A}'$ satisfies a strong $(\alpha_{A'}, \beta_{A'})$ -hash property. Let $p_{\hat{A}}$ be the joint distribution on $\hat{\mathcal{A}} \equiv \mathcal{A} \times \mathcal{A}'$ defined as

$$p_{\hat{A}}(A, A') \equiv p_A(A)p_{A'}(A')$$

for each $A \in \mathcal{A}$ and $A' \in \mathcal{A}'$. Let $(\alpha_{\hat{A}}, \beta_{\hat{A}})$ be defined as

$$\begin{aligned}\alpha_{\hat{A}} &\equiv \alpha_A \alpha_{A'} \\ \beta_{\hat{A}} &\equiv \beta_A + \beta_{A'}.\end{aligned}$$

Then the ensemble $(\hat{\mathcal{A}}, p_{\hat{A}})$ defined as

$$\begin{aligned}\hat{A} : \mathcal{U}^n &\rightarrow \text{Im}\mathcal{A} \times \text{Im}\mathcal{A}' \\ \hat{A}\mathbf{u} &\equiv (A\mathbf{u}, A'\mathbf{u})\end{aligned}$$

for each $\hat{A} \equiv (A, A') \in \hat{\mathcal{A}}$ and $\mathbf{u} \in \mathcal{U}^n$ has a strong $(\alpha_{\hat{A}}, \beta_{\hat{A}})$ -hash property.

In the following, we consider the combination of two ensembles, where the domains of functions are different. The following lemmas are essential for the proof of coding theorems presented in this paper. For a set \mathcal{A} of functions $A : \mathcal{U}^n \rightarrow \text{Im}\mathcal{A}$, a set \mathcal{B} of functions $B : \mathcal{V}^n \rightarrow \text{Im}\mathcal{B}$, let $p_{\mathbf{a}}$ and $p_{\mathbf{b}}$ the uniform distributions on $\text{Im}\mathcal{A}$ and $\text{Im}\mathcal{B}$, respectively, where random variables $\{A, B, \mathbf{a}, \mathbf{b}\}$ are mutually independent, that is,

$$p_{\text{ABab}}(A, B, \mathbf{a}, \mathbf{b}) = p_A(A)p_B(B)p_{\mathbf{a}}(\mathbf{a})p_{\mathbf{b}}(\mathbf{b})$$

for any A, B, \mathbf{a} , and \mathbf{b} .

Lemma 5: If (\mathcal{A}, p_A) and (\mathcal{B}, p_B) satisfy (H3), then

$$\begin{aligned}&p_{\text{AB}}(\{(A, B) : [\mathcal{G} \setminus \{(\mathbf{u}, \mathbf{v})\}] \cap [\mathcal{C}_A(A\mathbf{u}) \times \mathcal{C}_B(B\mathbf{v})] \neq \emptyset\}) \\ &\leq \frac{|\mathcal{G}|\alpha_A\alpha_B}{|\text{Im}\mathcal{A}||\text{Im}\mathcal{B}|} + \frac{\left[\max_{\mathbf{v} \in \mathcal{G}_{\mathcal{V}}} |\mathcal{G}_{\mathcal{U}|\mathcal{V}}(\mathbf{v})|\right] \alpha_A[\beta_B + 1]}{|\text{Im}\mathcal{A}|} + \frac{\left[\max_{\mathbf{u} \in \mathcal{G}_{\mathcal{U}}} |\mathcal{G}_{\mathcal{V}|\mathcal{U}}(\mathbf{u})|\right] \alpha_B[\beta_A + 1]}{|\text{Im}\mathcal{B}|} + \beta_A + \beta_B + \beta_A\beta_B\end{aligned}\quad (7)$$

for all $\mathcal{G} \subset \mathcal{U}^n \times \mathcal{V}^n$ and $(\mathbf{u}, \mathbf{v}) \in \mathcal{U}^n \times \mathcal{V}^n$.

Lemma 6: If (\mathcal{A}, p_A) and (\mathcal{B}, p_B) satisfy (H3), then

$$\begin{aligned}&p_{\text{ABab}}(\{(A, B, \mathbf{a}, \mathbf{b}) : \mathcal{T} \cap [\mathcal{C}_A(\mathbf{a}) \times \mathcal{C}_B(\mathbf{b})] = \emptyset\}) \\ &\leq \alpha_A\alpha_B - 1 + \frac{|\text{Im}\mathcal{B}| \left[\max_{\mathbf{v} \in \mathcal{T}_{\mathcal{V}}} |\mathcal{T}_{\mathcal{U}|\mathcal{V}}(\mathbf{v})|\right] \alpha_A[\beta_B + 1]}{|\mathcal{T}|} + \frac{|\text{Im}\mathcal{A}| \left[\max_{\mathbf{u} \in \mathcal{T}_{\mathcal{U}}} |\mathcal{T}_{\mathcal{V}|\mathcal{U}}(\mathbf{u})|\right] \alpha_B[\beta_A + 1]}{|\mathcal{T}|} \\ &\quad + \frac{|\text{Im}\mathcal{A}||\text{Im}\mathcal{B}|[\beta_A + \beta_B + \beta_A\beta_B + 1]}{|\mathcal{T}|}\end{aligned}\quad (8)$$

for all $\mathcal{T} \subset \mathcal{U}^n \times \mathcal{V}^n$.

It should be noted that Lemmas 5 and 6 are related to the collision-resistance property (Lemma 2) and the saturation property (Lemma 3), respectively. Proof are given in Appendix D.

B. Hash Property for Ensembles of Matrices

In the following, we discuss the hash property for an ensemble of matrices.

It has been discussed in the last section that the uniform distribution on the set of all linear functions has a strong $(1, 0)$ -hash property because it is a universal class of hash functions. In the following, we introduce another ensemble of matrices.

First, we introduce the average spectrum of an ensemble of matrices given in [1]. Let \mathcal{U} be a finite field and \mathcal{A} be a set of linear functions $A : \mathcal{U}^n \rightarrow \mathcal{U}^{l_A}$. It should be noted that A can be represented by a $l_A \times n$ matrix.

Let $\mathbf{t}(\mathbf{u})$ be the type² of $\mathbf{u} \in \mathcal{U}^n$, which is characterized by the empirical probability distribution $\nu_{\mathbf{u}}$ of the sequence \mathbf{u} . Let \mathcal{H} be a set of all types of length n except $\mathbf{t}(\mathbf{0})$, where $\mathbf{0}$ is the zero vector. For a probability distribution $p_{\mathbf{A}}$ on a set of $l_{\mathbf{A}} \times n$ matrices and a type \mathbf{t} , let $S(p_{\mathbf{A}}, \mathbf{t})$ be defined as

$$S(p_{\mathbf{A}}, \mathbf{t}) \equiv \sum_{A \in \mathcal{A}} p_{\mathbf{A}}(A) |\{\mathbf{u} \in \mathcal{U}^n : A\mathbf{u} = \mathbf{0}, \mathbf{t}(\mathbf{u}) = \mathbf{t}\}|,$$

which is called the expected number of codewords that have type \mathbf{t} in the context of linear codes. For given $\hat{\mathcal{H}}_{\mathbf{A}} \subset \mathcal{H}$, we define $\alpha_{\mathbf{A}}(n)$ and $\beta_{\mathbf{A}}(n)$ as

$$\alpha_{\mathbf{A}}(n) \equiv \frac{|\text{Im} \mathcal{A}|}{|\mathcal{U}|^{l_{\mathbf{A}}}} \cdot \max_{\mathbf{t} \in \hat{\mathcal{H}}_{\mathbf{A}}} \frac{S(p_{\mathbf{A}}, \mathbf{t})}{S(u_{\mathcal{A}}, \mathbf{t})} \quad (9)$$

$$\beta_{\mathbf{A}}(n) \equiv \sum_{\mathbf{t} \in \mathcal{H} \setminus \hat{\mathcal{H}}_{\mathbf{A}}} S(p_{\mathbf{A}}, \mathbf{t}), \quad (10)$$

where $u_{\mathcal{A}}$ denotes the uniform distribution on the set of all $l_{\mathbf{A}} \times n$ matrices.

We have the following theorem. The proof is given in Section VIII-A.

Theorem 1: Let $(\mathcal{A}, p_{\mathbf{A}})$ be an ensemble of matrices and assume that $p_{\mathbf{A}}(\{A : A\mathbf{u} = \mathbf{0}\})$ depends on \mathbf{u} only through the type $\mathbf{t}(\mathbf{u})$. If $(\alpha_{\mathbf{A}}, \beta_{\mathbf{A}})$, defined by (9) and (10), satisfies (H1) and (H2), then $(\mathcal{A}, p_{\mathbf{A}})$ has a strong $(\alpha_{\mathbf{A}}, \beta_{\mathbf{A}})$ -hash property.

Next, we introduce the ensemble of q -ary sparse matrices introduced in [17], which is the q -ary extension of the ensemble proposed in [14]. Let $\mathcal{U} \equiv \text{GF}(q)$ and $l_{\mathbf{A}} \equiv nR$ for given $0 < R < 1$. We generate an $l_{\mathbf{A}} \times n$ matrix A with the following procedure, where at most τ random nonzero elements are introduced in every row.

- 1) Start from an all-zero matrix.
- 2) For each $i \in \{1, \dots, n\}$, repeat the following procedure τ times:
 - a) Choose $(j, a) \in \{1, \dots, l_{\mathbf{A}}\} \times [\text{GF}(q) \setminus \{0\}]$ uniformly at random.
 - b) Add³ a to the (j, i) -element of A .

Assume that $\tau = O(\log n)$ is even and let $(\mathcal{A}, p_{\mathbf{A}})$ be an ensemble corresponding to the above procedure. Let $\hat{\mathcal{H}}_{\mathbf{A}} \subset \mathcal{H}$ be a set of types satisfying the requirement that the weight (the number of occurrences of non-zero elements) is large enough. Let $(\alpha_{\mathbf{A}}, \beta_{\mathbf{A}})$ be defined by (9) and (10). Then $\alpha_{\mathbf{A}}$ measures the difference between the ensemble $(\mathcal{A}, p_{\mathbf{A}})$ and the ensemble of all $l_{\mathbf{A}} \times n$ matrices with respect to the high-weight part of the average spectrum, and $\beta_{\mathbf{A}}$ provides the upper bound of the probability that the code $\{\mathbf{u} \in \mathcal{U}^n : A\mathbf{u} = \mathbf{0}\}$ has low-weight codewords. It is proved in [17, Theorem 2] that $(\alpha_{\mathbf{A}}, \beta_{\mathbf{A}})$ satisfy (H1) and (H2) by adopting an appropriate $\hat{\mathcal{H}}_{\mathbf{A}}$. Then, from Theorem 1, we have the fact that this ensemble has a strong $(\alpha_{\mathbf{A}}, \beta_{\mathbf{A}})$ -hash property. It should be noted that the convergence speed of $(\alpha_{\mathbf{A}}, \beta_{\mathbf{A}})$ depends on how fast τ grows in relation to the block length. The analysis of $(\alpha_{\mathbf{A}}, \beta_{\mathbf{A}})$ is given in the proof of [17, Theorem 2].

²In [17], it is called a histogram which is characterized by the number $n\nu_{\mathbf{u}}$ of occurrences of each symbol in the sequence \mathbf{u} . The type and the histogram is essentially the same.

³It should be noted that (j, i) -element of matrix is not overwritten by a when the same j is chosen again.

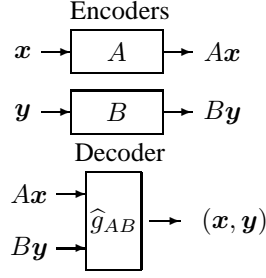


Fig. 3. Construction of Slepian-Wolf Source Code

IV. SLEPIAN-WOLF SOURCE CODING

In this section, we consider the Slepian-Wolf source coding illustrated in Fig. 1. The achievable rate region for this problem is given by

$$\begin{aligned} R_X &\geq H(X|Y) \\ R_Y &\geq H(Y|X) \\ R_X + R_Y &\geq H(X, Y), \end{aligned}$$

where (R_X, R_Y) denotes an encoding rate pair.

The achievability of the Slepian-Wolf source coding is proved in [4] and [5] for an ensemble of bin-coding and all q -ary linear matrices, respectively. The construction of encoders using sparse matrices is studied in [20][27] and the achievability is proved in [16] by using ML decoding. The aim of this section is to demonstrate the proof of the coding theorem based on the hash property. The proof is given in Section VIII-B.

We fix functions

$$\begin{aligned} A : \mathcal{X}^n &\rightarrow \text{Im}\mathcal{A} \\ B : \mathcal{Y}^n &\rightarrow \text{Im}\mathcal{B}, \end{aligned}$$

which are available for constructing encoders and a decoder. We define the encoders and the decoder (illustrated in Fig. 3)

$$\begin{aligned} \varphi_X : \mathcal{X}^n &\rightarrow \text{Im}\mathcal{A} \\ \varphi_Y : \mathcal{Y}^n &\rightarrow \text{Im}\mathcal{B} \\ \varphi^{-1} : \text{Im}\mathcal{A} \times \text{Im}\mathcal{B} &\rightarrow \mathcal{X}^n \times \mathcal{Y}^n \end{aligned}$$

as

$$\begin{aligned} \varphi_X(\mathbf{x}) &\equiv A\mathbf{x} \\ \varphi_Y(\mathbf{y}) &\equiv B\mathbf{y} \\ \varphi^{-1}(\mathbf{a}, \mathbf{b}) &\equiv \hat{g}_{AB}(\mathbf{a}, \mathbf{b}), \end{aligned}$$

where

$$\hat{g}_{AB}(\mathbf{a}, \mathbf{b}) \equiv \arg \min_{(\mathbf{x}', \mathbf{y}') \in \mathcal{C}_A(\mathbf{a}) \times \mathcal{C}_B(\mathbf{b})} D(\nu_{\mathbf{x}'\mathbf{y}'} \| \mu_{XY}).$$

It should be noted that the construction is analogous to the coset encoding/decoding in the case when A and B are linear functions.

The encoding rate pair (R_X, R_Y) is given by

$$R_X \equiv \frac{\log |\text{Im} \mathcal{A}|}{n}$$

$$R_Y \equiv \frac{\log |\text{Im} \mathcal{B}|}{n}$$

and the error probability $\text{Error}_{XY}(A, B)$ is given by

$$\text{Error}_{XY}(A, B) \equiv \mu_{XY}(\{(\mathbf{x}, \mathbf{y}) : \varphi^{-1}(\varphi_X(\mathbf{x}), \varphi_Y(\mathbf{y})) \neq (\mathbf{x}, \mathbf{y})\}).$$

We have the following theorem. It should be noted that \mathcal{X} and \mathcal{Y} are allowed to be non-binary and the correlation of the two sources is allowed to be asymmetric.

Theorem 2: Assume that (\mathcal{A}, p_A) and (\mathcal{B}, p_B) have a strong hash property. Let (X, Y) be a pair of stationary memoryless sources. If (R_X, R_Y) satisfies

$$R_X > H(X|Y) \tag{11}$$

$$R_Y > H(Y|X) \tag{12}$$

$$R_X + R_Y > H(X, Y), \tag{13}$$

then for any $\delta > 0$ and all sufficiently large n there are functions (sparse matrices) $A \in \mathcal{A}$ and $B \in \mathcal{B}$ such that

$$\text{Error}_{XY}(A, B) \leq \delta.$$

Remark 2: Instead of a maximum-likelihood decoder, we use a minimum-divergence decoder, which is compatible with typical-set decoding. We can replace the minimum-divergence decoder by

$$g'_{AB}(\mathbf{a}, \mathbf{b}) \equiv \arg \max_{\substack{(\mathbf{x}', \mathbf{y}') \in \mathcal{C}_A(\mathbf{a}) \times \mathcal{C}_B(\mathbf{b}) \\ \mathbf{x}' \in \mathcal{T}_{X,\gamma} \\ \mathbf{y}' \in \mathcal{T}_{Y,\gamma}}} \mu_{XY}(\mathbf{x}', \mathbf{y}'), \tag{14}$$

where $\gamma > 0$ should be defined properly. The reason for introducing the conditions $\mathbf{x}' \in \mathcal{T}_{X,\gamma}$ and $\mathbf{y}' \in \mathcal{T}_{Y,\gamma}$ is given at the end of Section VIII-B.

V. BROADCAST CHANNEL CODING

In this section we consider the broadcast channel coding problem illustrated in Fig. 2. A broadcast channel is characterized by the conditional probability distribution $\mu_{YZ|X}$, where X is a random variable corresponding to the channel input of a sender, and (Y, Z) is a pair of random variables corresponding to the channel outputs of respective receivers. In the following, we consider the case when a sender transmits two independent messages with nothing in common to two receivers. It is known that if a rate pair (R_Y, R_Z) satisfies

$$R_Y \leq I(U; Y) \tag{15}$$

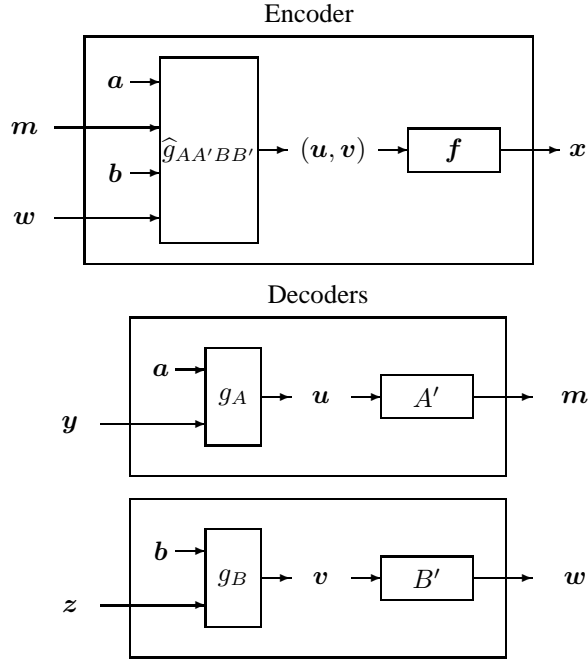


Fig. 4. Construction of Broadcast Channel Code

$$R_Z \leq I(V; Z) \quad (16)$$

$$R_Y + R_Z \leq I(U; Y) + I(V; Z) - I(U; V) \quad (17)$$

for some joint probability distribution μ_{UVX} on $\mathcal{U} \times \mathcal{V} \times \mathcal{X}$, then there is a code for this channel such that the decoding error probability goes to zero as the block length goes to infinity, where the joint distribution of random variable (U, V, X, Y, Z) is given by

$$\mu_{UVXYZ}(u, v, x, y, z) \equiv \mu_{YZ|X}(y, z|x) \mu_{UVX}(u, v, x). \quad (18)$$

This type of achievable region is derived in [15] and a simpler proof is given in [7]. Furthermore, it is shown in [10, Theorem 1] that it is sufficient to consider a joint probability distribution μ_{UVX} satisfying $|\mathcal{U}| \leq |\mathcal{X}|$, $|\mathcal{V}| \leq |\mathcal{X}|$, and $H(X|UV) = 0$. In the following, we assume that $|\mathcal{U}| \leq |\mathcal{X}|$, $|\mathcal{V}| \leq |\mathcal{X}|$, and there is a function $f : \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{X}$ such that

$$\mu_{UVX}(u, v, x) = \mu_{UV}(u, v) \chi(x = f(u, v)), \quad (19)$$

where (19) is equivalent to $H(X|UV) = 0$.

Let (R_Y, R_Z) be a pair of encoding rates. In the following, we assume that (R_Y, R_Z) , (r_Y, r_Z) , and ε satisfy

$$r_Y > H(U|Y) \quad (20)$$

$$r_Z > H(V|Z) \quad (21)$$

$$r_Y + R_Y < H(U) - \varepsilon \quad (22)$$

$$r_Z + R_Z < H(V) - \varepsilon \quad (23)$$

$$r_Y + R_Y + r_Z + R_Z < H(U, V) \quad (24)$$

$$r_Y + R_Y + r_Z + R_Z > H(U, V) - \varepsilon \quad (25)$$

for given $\mu_{YZ|X}$, μ_{UV} , and f . It should be noted that there are (r_Y, r_Z) and ε when (R_Y, R_Z) satisfies

$$R_Y < I(U; Y) \quad (26)$$

$$R_Z < I(V; Z) \quad (27)$$

$$R_Y + R_Z < I(U; Y) + I(V; Z) - I(U; V). \quad (28)$$

This fact is shown from the relations

$$H(U) - H(U|Y) = I(U; Y)$$

$$H(V) - H(V|Z) = I(V; Z)$$

$$\begin{aligned} H(U, V) - H(U|Y) - H(V|Z) &= I(U; Y) + I(V; Z) \\ &\quad - I(U; V). \end{aligned}$$

We fix functions

$$A : \mathcal{U}^n \rightarrow \text{Im}\mathcal{A}$$

$$A' : \mathcal{U}^n \rightarrow \text{Im}\mathcal{A}'$$

$$B : \mathcal{V}^n \rightarrow \text{Im}\mathcal{B}$$

$$B' : \mathcal{V}^n \rightarrow \text{Im}\mathcal{B}'$$

and vectors

$$\mathbf{a} \in \text{Im}\mathcal{A}$$

$$\mathbf{b} \in \text{Im}\mathcal{B}$$

available for an encoder and decoders satisfying

$$\begin{aligned} r_Y &= \frac{\log |\text{Im}\mathcal{A}|}{n} \\ R_Y &= \frac{\log |\text{Im}\mathcal{A}'|}{n} \\ r_Z &= \frac{\log |\text{Im}\mathcal{B}|}{n} \\ R_Z &= \frac{\log |\text{Im}\mathcal{B}'|}{n}. \end{aligned}$$

We define the encoder and the decoders

$$\varphi : \text{Im}\mathcal{A}' \times \text{Im}\mathcal{B}' \rightarrow \mathcal{X}^n$$

$$\varphi_Y^{-1} : \mathcal{Y}^n \rightarrow \text{Im}\mathcal{A}'$$

$$\varphi_Z^{-1} : \mathcal{Z}^n \rightarrow \text{Im}\mathcal{B}'$$

as

$$\varphi(\mathbf{m}, \mathbf{w}) \equiv \mathbf{f}(g_{AA'BB'}(\mathbf{a}, \mathbf{m}, \mathbf{b}, \mathbf{w}))$$

$$\begin{aligned}\varphi_Y^{-1}(\mathbf{y}) &\equiv A'g_A(\mathbf{a}|\mathbf{y}) \\ \varphi_Z^{-1}(\mathbf{z}) &\equiv B'g_B(\mathbf{b}|\mathbf{z}),\end{aligned}$$

where

$$\begin{aligned}\mathbf{f}(\mathbf{u}, \mathbf{v}) &\equiv (f(u_1, v_1), \dots, f(u_n, v_n)) \\ \hat{g}_{AA'BB'}(\mathbf{a}, \mathbf{m}, \mathbf{b}, \mathbf{w}) &\equiv \arg \min_{\substack{(\mathbf{u}', \mathbf{v}'): \\ \mathbf{u}' \in \mathcal{C}_A(\mathbf{a}) \cap \mathcal{C}_{A'}(\mathbf{m}) \\ \mathbf{v}' \in \mathcal{C}_B(\mathbf{b}) \cap \mathcal{C}_{B'}(\mathbf{w})}} D(\nu_{\mathbf{u}'\mathbf{v}'} \| \mu_{UV}) \\ g_A(\mathbf{a}|\mathbf{y}) &\equiv \arg \max_{\mathbf{u}' \in \mathcal{C}_A(\mathbf{a})} \mu_{U|Y}(\mathbf{u}'|\mathbf{y}) \\ g_B(\mathbf{b}|\mathbf{z}) &\equiv \arg \max_{\mathbf{v}' \in \mathcal{C}_B(\mathbf{b})} \mu_{V|Z}(\mathbf{v}'|\mathbf{z}).\end{aligned}$$

Let \mathbf{m} and \mathbf{w} be random variables corresponding to messages \mathbf{m} and \mathbf{w} , respectively, where the probability distributions $p_{\mathbf{m}}$ and $p_{\mathbf{w}}$ are given by

$$p_{\mathbf{m}}(\mathbf{m}) \equiv \begin{cases} \frac{1}{|\text{Im}\mathcal{A}'|} & \text{if } \mathbf{m} \in \text{Im}\mathcal{A}' \\ 0, & \text{if } \mathbf{m} \notin \text{Im}\mathcal{A}' \end{cases} \quad (29)$$

$$p_{\mathbf{w}}(\mathbf{w}) \equiv \begin{cases} \frac{1}{|\text{Im}\mathcal{B}'|} & \text{if } \mathbf{w} \in \text{Im}\mathcal{B}' \\ 0, & \text{if } \mathbf{w} \notin \text{Im}\mathcal{B}' \end{cases} \quad (30)$$

and the joint distribution $p_{\mathbf{m}\mathbf{w}YZ}$ of the messages and outputs is given by

$$p_{\mathbf{m}\mathbf{w}YZ}(\mathbf{m}, \mathbf{w}, \mathbf{y}, \mathbf{z}) \equiv \mu_{YZ|X}(\mathbf{y}, \mathbf{z} | \varphi(\mathbf{m}, \mathbf{w})) p_M(\mathbf{m}) p_W(\mathbf{w}).$$

Let us remark an intuitive interpretation of the code construction, which is illustrated in Fig. 4. Assume that \mathbf{a} and \mathbf{b} are shared by the encoder and the decoder. For \mathbf{a} , \mathbf{b} , and messages \mathbf{m} , \mathbf{w} , the function $\hat{g}_{AA'BB'}$ generates a pair of typical sequence $(\mathbf{u}, \mathbf{v}) \in \mathcal{T}_{UV, \gamma}$ and it is converted to a channel input \mathbf{x} by using a function f , where $(\mathbf{u}, \mathbf{v}, \mathbf{x})$ is jointly-typical. One decoder reproduces \mathbf{u} by using g_A from \mathbf{a} and a channel output \mathbf{y} and another decoder reproduces \mathbf{v} by using g_B from \mathbf{b} and a channel output \mathbf{z} . Since $A'\mathbf{u} = \mathbf{m}$, $B'\mathbf{v} = \mathbf{w}$, and $(\mathbf{u}, \mathbf{v}, \mathbf{x}, \mathbf{y})$ is jointly typical, the decoding succeeds if the amount of information of \mathbf{a} (resp. \mathbf{b}) is greater than $H(U|Y)$ (resp. $H(V|Z)$) to satisfy the collision-resistance property. On the other hand, the rate of \mathbf{a} , \mathbf{m} , \mathbf{b} , and \mathbf{w} should satisfy the saturation property to find a jointly typical sequences (\mathbf{u}, \mathbf{v}) with probability close to 1, that is, the right hand side of (8) should tend to zero as the block length goes to infinity. Then we can set the encoding rate of \mathbf{m} and \mathbf{w} satisfying (26)–(28).

Let $\text{Error}_{YZ|X}(A, A', B, B', \mathbf{a}, \mathbf{b})$ be the decoding error probability given by

$$\begin{aligned}\text{Error}_{YZ|X}(A, A', B, B', \mathbf{a}, \mathbf{b}) \\ \equiv 1 - \sum_{\mathbf{m}, \mathbf{w}, \mathbf{y}, \mathbf{z}} \mu_{YZ|X}(\mathbf{y}, \mathbf{z} | \varphi(\mathbf{m}, \mathbf{w})) p_M(\mathbf{m}) p_W(\mathbf{w}) \chi(\varphi_Y^{-1}(\mathbf{y}) = \mathbf{m}) \chi(\varphi_Z^{-1}(\mathbf{z}) = \mathbf{w}).\end{aligned} \quad (31)$$

Then we have the following theorem. It should be noted that the channel is allowed to be asymmetric and non-degraded.

Theorem 3: Let $\mu_{YZ|X}$ be the conditional probability distribution of a stationary memoryless channel and μ_{UVXYZ} be defined by (18) and (19) for a given joint probability distribution μ_{UV} and a function f . For a

given (R_Y, R_Z) , (r_Y, r_Z) , and ε satisfying (20)–(25), assume that ensembles (\mathcal{A}, p_A) , $(\mathcal{A}', p_{A'})$, (\mathcal{B}, p_B) , and $(\mathcal{B}', p_{B'})$ have a strong hash property. Then, for any $\delta > 0$ and all sufficiently large n , there are functions (sparse matrices) $A \in \mathcal{A}$, $A' \in \mathcal{A}'$, $B \in \mathcal{B}$, $B' \in \mathcal{B}'$ and vectors $\mathbf{a} \in \text{Im}\mathcal{A}$, $\mathbf{b} \in \text{Im}\mathcal{B}$ such that

$$\text{Rate}_Y > R_Y - \delta \quad (32)$$

$$\text{Rate}_Z > R_Z - \delta \quad (33)$$

$$\text{Error}_{YZ|X}(A, A', B, B', \mathbf{a}, \mathbf{b}) < \delta \quad (34)$$

By assuming that $\delta \rightarrow 0$, the rate of the proposed code is close to the boundary of the region specified by (15)–(17) for a given μ_{UV} and f . ■

Remark 3: It should be noted that the maximum-likelihood decoders g_A and g_B can be replaced by the minimum-divergence decoders \hat{g}_A and \hat{g}_B defined as

$$\hat{g}_A(\mathbf{a}|\mathbf{y}) \equiv \arg \min_{\mathbf{u}' \in \mathcal{C}_A(\mathbf{a})} D(\nu_{\mathbf{u}'|\mathbf{y}} \| \mu_{U|Y}|\nu_{\mathbf{y}})$$

$$\hat{g}_B(\mathbf{b}|\mathbf{z}) \equiv \arg \min_{\mathbf{v}' \in \mathcal{C}_B(\mathbf{b})} D(\nu_{\mathbf{v}'|\mathbf{z}} \| \mu_{V|Z}|\nu_{\mathbf{z}}),$$

respectively.

Remark 4: In [7][15], the set of strong typical sequences are used for the construction of a broadcast channel code, where there is no structure of the set of codewords. It should be noted that the encoder and the decoder have to share the large (exponentially in the block length) table which indicates the correspondence between an index and codewords. The time complexity of the decoding grows exponentially in the block length. They represent obstacles to the implementation. In our construction, almost all codewords have strong typicality but the set of codewords have a structure specified by functions A , A' , B , B' and vectors \mathbf{a} , \mathbf{b} . When the functions are linear, it is expected that the space and time complexity can be reduced compared with conventional constructions.

Remark 5: It should be noted that the inner bound presented in the beginning of this section is not convex in general and we can apply the time sharing principle to obtain the convex hull of this region. It should also be noted that, according to [8, Page 9-36], there is a much larger inner region given by the set of a rate pair (R_Y, R_Z) satisfying

$$R_Y \leq I(S, U; Y)$$

$$R_Z \leq I(S, V; Z)$$

$$R_Y + R_Z \leq I(S, U; Y) + I(V; Z|S) - I(U; V|S)$$

$$R_Y + R_Z \leq I(U; Y|S) + I(S, V; Z) - I(U; V|S),$$

for some probability distribution μ_{SUV} on $\mathcal{S} \times \mathcal{U} \times \mathcal{V}$ and a function $f : \mathcal{S} \times \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{X}$. We presented the code construction for the smaller inner region. because it is a typical application of the extended saturation property (Lemma 6). It is a future challenge to obtain a code construction for the larger inner region based on the strong hash property.

VI. CONSTRUCTION OF MINIMUM-DIVERGENCE OPERATION

In this section, we introduce the construction of the minimum-divergence operation

$$\hat{g}_{\hat{A}\hat{B}}(\hat{\mathbf{a}}, \hat{\mathbf{b}}) \equiv \arg \min_{(\mathbf{u}', \mathbf{v}') \in \mathcal{C}_{\hat{A}}(\hat{\mathbf{a}}) \times \mathcal{C}_{\hat{B}}(\hat{\mathbf{b}})} D(\nu_{\mathbf{u}'\mathbf{v}'} \| \mu_{UV}),$$

which is used to construct the encoder of the broadcast channel and the decoder of the Slepian-Wolf codes.

The construction of this operation is introduced in [18] for a binary alphabet. However, we have to construct this operation for an alphabet that is a product space $\mathcal{U} \times \mathcal{V}$.

In the following, we assume that $\{0, 1\}$ is a finite field, $\mathcal{U} = \mathcal{V} = \{0, 1\}$ and that \hat{A} and \hat{B} are matrices on $\{0, 1\}$. It should be noted that our construction of broadcast channel codes can approach the boundary of the achievable region by assuming $|\mathcal{X}| = 2$ (see [10, Theorem 1]).

Let $t_{\mathbf{uv}}(u, v) \in \{0, \dots, n\}$ be the number of occurrences of $(u, v) \in \{0, 1\}^2$ in $(\mathbf{u}, \mathbf{v}) \in \{0, 1\}^n \times \{0, 1\}^n$. For example, when $n = 8$ and

$$\begin{aligned} \mathbf{u} &= 01001010 \\ \mathbf{v} &= 00101001, \end{aligned}$$

we have

$$\begin{aligned} t_{\mathbf{uv}}(0, 0) &= 3 \\ t_{\mathbf{uv}}(0, 1) &= 2 \\ t_{\mathbf{uv}}(1, 0) &= 2 \\ t_{\mathbf{uv}}(1, 1) &= 1. \end{aligned}$$

Let $\mathbf{t}_{\mathbf{uv}} \equiv \{t_{\mathbf{uv}}(0, 0), t_{\mathbf{uv}}(0, 1), t_{\mathbf{uv}}(1, 0), t_{\mathbf{uv}}(1, 1)\}$. We have

$$\hat{g}_{\hat{A}\hat{B}}(\hat{\mathbf{a}}, \hat{\mathbf{b}}) = \arg \min_{\substack{(\mathbf{u}_t, \mathbf{v}_t): \\ \mathbf{t} \in \mathcal{H}}} D(\nu_{\mathbf{u}_t\mathbf{v}_t} \| \mu_{UV}),$$

where

$$\begin{aligned} (\mathbf{u}_t, \mathbf{v}_t) &\equiv \begin{cases} (\mathbf{u}, \mathbf{v}) & \text{if } \exists (\mathbf{u}, \mathbf{v}) \text{ s.t. } \mathbf{t}_{\mathbf{uv}} = \mathbf{t}, \\ & \hat{A}\mathbf{u} = \hat{\mathbf{a}}, \hat{B}\mathbf{v} = \hat{\mathbf{b}} \\ \text{'error'} & \text{otherwise} \end{cases} \\ \mathcal{H} &\equiv \left\{ \mathbf{t} : t(u, v) \in \{0, \dots, n\}, \sum_{u, v} t(u, v) = n \right\} \end{aligned} \tag{35}$$

and $D(\nu_{\mathbf{u}_t\mathbf{v}_t} \| \mu_{UV}) \equiv \infty$ when the operation (35) returns an error. It should be noted that the cardinality of the set \mathcal{H} is at most $[n+1]^4$.

We can apply the linear programming technique introduced in [9] to the construction of (35). By using the technique introduced in [9], the conditions $\hat{A}\mathbf{u} = \hat{\mathbf{a}}$ and $\hat{B}\mathbf{v} = \hat{\mathbf{b}}$ can be replaced by linear inequalities. In the following, we focus on the condition $\mathbf{t}_{\mathbf{uv}} = \mathbf{t}$ for a given $\mathbf{t} \in \mathcal{H}$.

Let $\mathbf{u} \equiv (u_1, \dots, u_n)$, $\mathbf{v} \equiv (v_1, \dots, v_n)$ and $\mathbf{t} \equiv (t(0,0), t(0,1), t(1,0), t(1,1))$. First, we introduce the auxiliary variable $\mathbf{s}(u, v) \equiv (s_1(u, v), \dots, s_n(u, v))$ for $(u, v) \in \{0, 1\}^2$ defined as

$$s_i(u, v) \equiv \begin{cases} 1, & \text{if } u_i = u, v_i = v \\ 0, & \text{otherwise.} \end{cases} \quad (36)$$

Then the condition $\mathbf{t}_{uv} = \mathbf{t}$ is equivalent to linear equalities

$$\sum_{i=1}^n s_i(u, v) = t(u, v)$$

for all $(u, v) \in \{0, 1\}^2$.

Next, we show that the relation (36) can be replaced by linear inequalities. Let $\mathcal{S}(u, v)$ be defined as

$$\mathcal{S}(u, v) \equiv \{(1, u, v)\} \cup \left[\bigcup_{(u', v') \neq (u, v)} \{(0, u', v')\} \right].$$

Then we obtain the convex hull of $\mathcal{S}(u, v)$ defined by the set of (s', u', v') that satisfies the following linear inequalities.

$$\begin{aligned} s' &\geq 0 \\ u' &\geq 0 \quad \text{if } u = 0 \\ u' &\leq 1 \quad \text{if } u = 1 \\ v' &\geq 0 \quad \text{if } v = 0 \\ v' &\leq 1 \quad \text{if } v = 1 \\ s' + [-1]^u u' &\leq 1 - u \\ s' + [-1]^v v' &\leq 1 - v \\ s' + [-1]^u u' + [-1]^v v' &\geq 1 - u - v, \end{aligned}$$

where constants u, v and variables s', u', v' are considered to be real numbers. For example, when $(u, v) = (0, 0)$, we have

$$\mathcal{S}(0, 0) = \{(1, 0, 0), (0, 0, 1), (0, 1, 0), (0, 1, 1)\}$$

and linear inequalities

$$\begin{aligned} s' &\geq 0 \\ u' &\geq 0 \\ v' &\geq 0 \\ s' + u' &\leq 1 \\ s' + v' &\leq 1 \\ s' + u' + v' &\geq 1. \end{aligned}$$

Finally, we have $24n + 4$ inequalities, which replace the condition $\mathbf{t}_{uv} = \mathbf{t}$, defined as

$$\sum_{i=1}^n s_i(u, v) = t(u, v)$$

$$s_i(u, v) \geq 0$$

$$u_i \geq 0 \quad \text{if } u = 0$$

$$u_i \leq 1 \quad \text{if } u = 1$$

$$v_i \geq 0 \quad \text{if } v = 0$$

$$v_i \leq 1 \quad \text{if } v = 1$$

$$s_i(u, v) + [-1]^u u_i \leq 1 - u$$

$$s_i(u, v) + [-1]^v v_i \leq 1 - v$$

$$s_i(u, v) + [-1]^u u_i + [-1]^v v_i \geq 1 - u - v,$$

where we take all $i \in \{1, 2, \dots, n\}$ and $(u, v) \in \{0, 1\}^2$.

Remark 6: It should be noted here that linear programming frequently finds non-integral solutions for the operation (35) even if \hat{A} and \hat{B} are sparse matrices. Let us consider the polytope with vertexes $\mathcal{C}_{\hat{A}}(\hat{\mathbf{a}}) \times \mathcal{C}_{\hat{B}}(\hat{\mathbf{b}})$. Then the section of the polytope cut by the hyperplane defined by $\mathbf{t}_{uv} = \mathbf{t}$ may have many fractional points. This is one reason why the linear programming finds non-integral solutions. In fact, the original operation introduced in [18] has the same problem. It remains a future challenge to find a good implementation of the operation (35).

VII. EXTENSION TO THREE OR MORE TERMINALS

In this section, we extend our results to three or more terminals. We use the following notations:

$$\mathcal{K} \equiv \{1, 2, \dots, k\}$$

$$\mathcal{U}_{\mathcal{K}} \equiv \times_{j \in \mathcal{K}} \mathcal{U}_j$$

$$\mathbf{u}_{\mathcal{K}} \equiv \{(u_{j,1}, u_{j,2}, \dots, u_{j,n})\}_{j \in \mathcal{K}}$$

$$A_{\mathcal{K}} \equiv \{A_j\}_{j \in \mathcal{K}}$$

and

$$U_{\mathcal{J}} \equiv \{U_j\}_{j \in \mathcal{J}}$$

for each $\mathcal{J} \subset \mathcal{K}$.

A. Multiple Slepian-Wolf Source Code

In this section, we consider the Slepian-Wolf coding for k correlated sources.

Let $X_{\mathcal{K}} \in \mathcal{X}_{\mathcal{K}}$ be correlated sources and $\mu_{X_{\mathcal{K}}}$ be the joint probability distribution of $X_{\mathcal{K}}$. For each $j \in \mathcal{K}$, we consider an ensemble $(\mathcal{A}_j, \mathbf{p}_{A_j})$ of functions $A_j : \mathcal{X}_j^n \rightarrow \text{Im} \mathcal{A}_j$. We assume that function A_j is shared by the j -th encoder and the decoder. We define the j -th encoder as

$$\varphi_j(\mathbf{x}_j) \equiv A_j \mathbf{x}_j$$

for each output $\mathbf{x}_j \in \mathcal{X}_j^n$ of the j -th source. Let \widehat{g}_{A_K} be defined as

$$\widehat{g}_{A_K}(\mathbf{a}_K) \equiv \arg \min_{\mathbf{x}_K \in \mathcal{C}_{A_K}(\mathbf{a}_K)} D(\nu_{\mathbf{x}_K} \parallel \mu_{X_K}),$$

where \mathbf{a}_j is the codeword of the j -th encoder and

$$\mathcal{C}_{A_K}(\mathbf{a}_K) \equiv \{\mathbf{u}_K : A_j \mathbf{u}_j = \mathbf{a}_j \text{ for all } j \in K\}.$$

Then we define the decoder as

$$\varphi^{-1}(\mathbf{a}_K) \equiv \widehat{g}_{A_K}(\mathbf{a}_K).$$

The construction of the above minimum-divergence decoder is described in Section VII-D.

For each $j \in K$, the encoding rate R_j for the j -th encoder is given by

$$R_j \equiv \frac{\log |\text{Im} \mathcal{A}_j|}{n}$$

and the error probability is given by

$$\text{Error}_X(A_K) \equiv \mu_{X_K} \left(\left\{ \mathbf{x}_K : \varphi^{-1} \left(\{\varphi_j(\mathbf{x}_j)\}_{j \in K} \right) \neq \mathbf{x}^K \right\} \right).$$

We have the following theorem.

Theorem 4: Assume that $\{R_j\}_{j \in K}$ satisfies

$$\sum_{j \in \mathcal{J}} R_j > H(X_{\mathcal{J}} | X_{\mathcal{J}^c})$$

for all $\mathcal{J} \subset K$ except the empty set, and an ensemble $(\mathcal{A}_j, \mathbf{p}_{A_j})$ has an $(\alpha_{A_j}, \beta_{A_j})$ -hash property for all $j \in K$.

Then for any $\delta > 0$ and all sufficiently large n there are functions (sparse matrices) $A_K \equiv \{A_j\}_{j \in K}$ such that

$$\text{Error}_X(A_K) \leq \delta.$$

B. Multiple Output Broadcast Channel Code

In this section, we consider the k -output broadcast channel coding.

Let $\mu_{Y_K|X}$ be the conditional distribution of a broadcast channel with input $X \in \mathcal{X}$ and k outputs $Y_K \in \mathcal{Y}_K$.

Let $U_K \in \mathcal{U}_K$ be a multiple random variable and μ_{U_K} be the joint probability distribution of U_K . Let $F : \mathcal{U}_K \rightarrow \mathcal{X}$ be a function that is allowed to be non-deterministic. The joint distribution $\mu_{U_K X Y_K}$ of (U_K, X, Y_K) is given by

$$\mu_{U_K X Y_K}(u_K, x, y_K) \equiv \mu_{Y_K|X}(y_K|x) \mu_{X|U_K}(x|u_K) \mu_{U_K}(u_K),$$

where $\mu_{X|U_K}$ represents the transition probability distribution of the function F .

For each $j \in K$, we consider an ensemble $(\mathcal{A}_j, \mathbf{p}_{A_j})$ of functions $A_j : \mathcal{U}_j^n \rightarrow \text{Im} \mathcal{A}_j$, where \mathcal{U}_j is an alphabet of the source U_j . We also consider an ensemble $(\mathcal{A}'_j, \mathbf{p}_{A'_j})$ of functions $A'_j : \mathcal{U}_j^n \rightarrow \text{Im} \mathcal{A}'_j$. Let \mathbf{m}_j be the random variable corresponding to the j -th message, where the probability distribution $p_{\mathbf{m}_j}$ are given by

$$p_{\mathbf{m}_j}(\mathbf{m}_j) \equiv \begin{cases} \frac{1}{|\text{Im} \mathcal{A}'_j|} & \text{if } \mathbf{m}_j \in \text{Im} \mathcal{A}'_j \\ 0, & \text{if } \mathbf{m}_j \notin \text{Im} \mathcal{A}'_j \end{cases}$$

for each $j \in \mathcal{K}$. We assume that functions A_j , A'_j and a vector $\mathbf{a}_j \in \text{Im} \mathcal{A}_j$ are shared by the encoder and the j -th decoder. For a multiple message $\mathbf{m}_{\mathcal{K}} \in \times_{j \in \mathcal{K}} \text{Im} \mathcal{A}'_j$, we define the encoder $\varphi : \times_{j \in \mathcal{K}} \text{Im} \mathcal{A}'_j \rightarrow \mathcal{X}^n$ as

$$\hat{g}_{A_{\mathcal{K}} A'_{\mathcal{K}}}(\mathbf{m}_{\mathcal{K}}, \mathbf{a}_{\mathcal{K}}) \equiv \arg \min_{\mathbf{u}_{\mathcal{K}} \in \mathcal{C}_{A_{\mathcal{K}} A'_{\mathcal{K}}}(\mathbf{m}_{\mathcal{K}}, \mathbf{a}_{\mathcal{K}})} D(\nu_{\mathbf{u}_{\mathcal{K}}} \| \mu_{U_{\mathcal{K}}})$$

$$\mathbf{F}(\mathbf{u}_{\mathcal{K}}) \equiv (F(u_{\mathcal{K},1}), \dots, F(u_{\mathcal{K},n}))$$

$$\varphi(\mathbf{m}_{\mathcal{K}}) \equiv \mathbf{F}(\hat{g}_{A_{\mathcal{K}} A'_{\mathcal{K}}}(\mathbf{m}_{\mathcal{K}}, \mathbf{a}_{\mathcal{K}})),$$

where $u_{\mathcal{K},i} \equiv \{u_{j,i}\}_{j \in \mathcal{K}}$ for each $i \in \{1, \dots, n\}$ and

$$\mathcal{C}_{A_{\mathcal{K}} A'_{\mathcal{K}}}(\mathbf{m}_{\mathcal{K}}, \mathbf{a}_{\mathcal{K}}) \equiv \left\{ \mathbf{u}_{\mathcal{K}} : \begin{array}{l} A_j \mathbf{u}_j = \mathbf{a}_j, A'_j \mathbf{u}_j = \mathbf{m}_j \\ \text{for all } j \in \mathcal{K} \end{array} \right\}.$$

The construction of the above minimum-divergence decoder is described in Section VII-D. We define the j -th decoder $\varphi_j^{-1} : \mathcal{Y}_j^n \rightarrow \text{Im} \mathcal{A}'_j$ as

$$g_{A_j}(\mathbf{a}_j | \mathbf{y}_j) \equiv \arg \max_{\mathbf{u}_j \in \mathcal{C}_{A_j}(\mathbf{a}_j)} \mu_{U_j | Y_j}(\mathbf{u}_j | \mathbf{y}_j)$$

$$\varphi_j^{-1}(\mathbf{y}_j) \equiv A'_j g_{A_j}(\mathbf{a}_j | \mathbf{y}_j),$$

for each output $\mathbf{y}_j \in \mathcal{Y}_j^n$. The error probability is given by

$$\text{Error}_{Y_{\mathcal{K}}|X}(A_{\mathcal{K}}, A'_{\mathcal{K}}, \mathbf{a}_{\mathcal{K}}) \equiv 1 - \sum_{\mathbf{y}_{\mathcal{K}}} \mu_{Y_{\mathcal{K}}|X}(\mathbf{y}_{\mathcal{K}} | \varphi(\mathbf{m}_{\mathcal{K}})) \prod_{j \in \mathcal{K}} [p_{\mathbf{m}_j}(\mathbf{m}_j) \chi(\varphi_j^{-1}(\mathbf{y}_j) = \mathbf{m}_j)].$$

We have the following theorem.

Theorem 5: For each $j \in \mathcal{K}$, let R_j be the encoding rate for the j -th encoder. Assume that $\{(r_j, R_j, \varepsilon_j)\}_{j \in \mathcal{K}}$ and ε satisfy

$$r_j = \frac{\log |\text{Im} \mathcal{A}_j|}{n} \tag{37}$$

$$R_j = \frac{\log |\text{Im} \mathcal{A}'_j|}{n} \tag{38}$$

$$r_j > H(U_j | Y_j) \tag{39}$$

$$\sum_{j \in \mathcal{J}} [R_j + r_j] < H(U_{\mathcal{J}}) - \varepsilon \tag{40}$$

$$H(U_{\mathcal{K}}) - \varepsilon < \sum_{j \in \mathcal{K}} [R_j + r_j] < H(U_{\mathcal{K}}) \tag{41}$$

for all non-empty set $\mathcal{J} \subsetneq \mathcal{K}$ and $j \in \mathcal{K}$. Furthermore, assume that an ensemble $(\mathcal{A}_j, \mathbf{p}_{A_j})$ (resp. $(\mathcal{A}'_j, \mathbf{p}_{A'_j})$) has an $(\alpha_{A_j}, \beta_{A_j})$ -hash (resp. $(\alpha_{A'_j}, \beta_{A'_j})$ -hash) property for all $j \in \mathcal{K}$. Then for any $\delta > 0$ and all sufficiently large n there are functions (sparse matrices) and vectors $(A_{\mathcal{K}}, A'_{\mathcal{K}}, \mathbf{a}_{\mathcal{K}}) \equiv \{(A_j, A'_j, \mathbf{a}_j)\}_{j \in \mathcal{K}}$ such that

$$\text{Error}_{Y_{\mathcal{K}}|X}(A_{\mathcal{K}}, A'_{\mathcal{K}}, \mathbf{a}_{\mathcal{K}}) \leq \delta.$$

It should be noted that there are $\{r_j\}_{j \in \mathcal{K}}$ and ε satisfying (37)–(41) if $\{R_j\}_{j \in \mathcal{K}}$ satisfy

$$\begin{aligned} \sum_{j \in \mathcal{J}} R_j &< H(U_{\mathcal{J}}) - \sum_{j \in \mathcal{J}} H(U_j | Y_j) \\ &= \sum_{j \in \mathcal{J}} I(U_j; Y_j) - \left[\sum_{j \in \mathcal{J}} H(U_j) - H(U_{\mathcal{J}}) \right] \end{aligned} \tag{42}$$

for all non-empty set $\mathcal{J} \subset \mathcal{K}$. The condition (42) is equivalent to the conditions

$$\begin{aligned} R_j &< I(U_j; Y_j) \text{ for } j \in \{1, 2, 3\} \\ R_1 + R_2 &< I(U_1; Y_1) + I(U_2; Y_2) - I(U_1; U_2) \\ R_1 + R_3 &< I(U_1; Y_1) + I(U_3; Y_3) - I(U_1; U_3) \\ R_2 + R_3 &< I(U_2; Y_2) + I(U_3; Y_3) - I(U_2; U_3) \\ R_1 + R_2 + R_3 &< I(U_1; Y_1) + I(U_2; Y_2) + I(U_3; Y_3) - I(U_1; U_2) - I(U_1, U_2; U_3) \end{aligned}$$

specified in [8, Page 9-47], which considers the case of $k = 3$.

C. Fundamental Lemmas for Theorems 4 and 5

We can prove Theorems 4 and 5 in a similar way to Theorems 2 and 3 by using the following lemmas proved in Appendix E. In the following, we define

$$|\text{Im}\mathcal{A}_{\mathcal{J}}| \equiv \prod_{j \in \mathcal{J}} |\text{Im}\mathcal{A}_j| \quad (43)$$

$$|\mathcal{T}_{\mathcal{J}|\mathcal{J}^c}| \equiv \begin{cases} 1, & \text{if } \mathcal{J} = \emptyset \\ |\mathcal{T}|, & \text{if } \mathcal{J} = \mathcal{K} \\ \max_{\mathbf{u}_{\mathcal{J}^c} \in \mathcal{T}_{\mathcal{U}_{\mathcal{J}^c}}} |\mathcal{T}_{\mathcal{U}_{\mathcal{J}}|\mathcal{U}_{\mathcal{J}^c}}(\mathbf{u}_{\mathcal{J}^c})|, & \text{otherwise.} \end{cases} \quad (44)$$

Lemma 7: For each $j \in \mathcal{K}$, let \mathcal{A}_j be a set of functions $A_j : \mathcal{U}_j^n \rightarrow \text{Im}\mathcal{A}_j$ and p_{A_j} be the probability distribution on \mathcal{A}_j , where (\mathcal{A}_j, p_{A_j}) satisfies (H3). Let the joint distribution $p_{\mathbf{A}_{\mathcal{K}}\mathbf{a}_{\mathcal{K}}}$ be defined as

$$p_{\mathbf{A}_{\mathcal{K}}\mathbf{a}_{\mathcal{K}}}(A_{\mathcal{K}}, \mathbf{a}_{\mathcal{K}}) \equiv \prod_{j \in \mathcal{K}} p_{A_j}(A_j) p_{\mathbf{a}_j}(\mathbf{a}_j)$$

for each $\{A_j, \mathbf{a}_j\}_{j \in \mathcal{K}}$. For each $\mathcal{J} \subset \mathcal{K}$, let $\alpha_{\mathbf{A}_{\mathcal{J}}}$ and $\beta_{\mathbf{A}_{\mathcal{J}}}$ be defined as

$$\alpha_{\mathbf{A}_{\mathcal{J}}} \equiv \prod_{j \in \mathcal{J}} \alpha_{A_j} \quad (45)$$

$$\beta_{\mathbf{A}_{\mathcal{J}}} \equiv \prod_{j \in \mathcal{J}} [\beta_{A_j} + 1] - 1. \quad (46)$$

Then

$$p_{\mathbf{A}_{\mathcal{K}}}(\{A_{\mathcal{K}} : [\mathcal{G} \setminus \{\mathbf{u}_{\mathcal{K}}\}] \cap \mathcal{C}_{A_{\mathcal{K}}}(A_{\mathcal{K}}\mathbf{u}_{\mathcal{K}}) \neq \emptyset\}) \leq \sum_{\substack{\mathcal{J} \subset \mathcal{K} \\ \mathcal{J} \neq \emptyset}} \frac{|\mathcal{G}_{\mathcal{J}|\mathcal{J}^c}| \alpha_{\mathbf{A}_{\mathcal{J}}} [\beta_{\mathbf{A}_{\mathcal{J}^c}} + 1]}{|\text{Im}\mathcal{A}_{\mathcal{J}}|} + \beta_{\mathbf{A}_{\mathcal{K}}} \quad (47)$$

for all $\mathcal{G} \subset [\mathcal{U}_{\mathcal{K}}]^n$ and $\mathbf{u}_{\mathcal{K}} \in [\mathcal{U}_{\mathcal{K}}]^n$, and

$$p_{\mathbf{A}_{\mathcal{K}}\mathbf{a}_{\mathcal{K}}}(\{(A_{\mathcal{K}}, \mathbf{a}_{\mathcal{K}}) : \mathcal{T} \cap \mathcal{C}_{A_{\mathcal{K}}}(\mathbf{a}_{\mathcal{K}}) = \emptyset\}) \leq \alpha_{\mathbf{A}_{\mathcal{K}}} - 1 + \sum_{\mathcal{J} \subsetneq \mathcal{K}} \frac{|\text{Im}\mathcal{A}_{\mathcal{J}^c}| |\mathcal{T}_{\mathcal{J}|\mathcal{J}^c}| \alpha_{\mathbf{A}_{\mathcal{J}}} [\beta_{\mathbf{A}_{\mathcal{J}^c}} + 1]}{|\mathcal{T}|} \quad (48)$$

for all $\mathcal{T} \subset [\mathcal{U}_{\mathcal{K}}]^n$. Furthermore, if $(\alpha_{\mathbf{A}_j}, \beta_{\mathbf{A}_j})$ satisfies (H1) and (H2) for all $j \in \mathcal{K}$, then

$$\lim_{n \rightarrow \infty} \alpha_{\mathbf{A}_{\mathcal{J}}}(n) = 1 \quad (49)$$

$$\lim_{n \rightarrow \infty} \beta_{\mathbf{A}_{\mathcal{J}}}(n) = 0 \quad (50)$$

for every $\mathcal{J} \subset \mathcal{K}$.

Lemma 8: If \mathcal{T} is a subset of $\mathcal{T}_{U_{\mathcal{K}}, \gamma}$, then

$$\max_{\mathbf{u}_{\mathcal{J}^c} \in \mathcal{T}_{U_{\mathcal{J}^c}}} |\mathcal{T}_{U_{\mathcal{J}}|U_{\mathcal{J}^c}}(\mathbf{u}_{\mathcal{J}^c})| \leq 2^n [H(U_{\mathcal{J}}|U_{\mathcal{J}^c}) + \eta_{U_{\mathcal{J}}|U_{\mathcal{J}^c}}(\gamma|\gamma)]$$

for every non-empty set $\mathcal{J} \subsetneq \mathcal{K}$.

D. Construction of Minimum-Divergence Operation

In this section, we introduce the construction of the minimum-divergence operation

$$\hat{g}_{A_{\mathcal{K}}}(\mathbf{a}_{\mathcal{K}}) \equiv \arg \min_{\mathbf{u}_{\mathcal{K}} \in \mathcal{C}_{A_{\mathcal{K}}}(\mathbf{a}_{\mathcal{K}})} D(\nu_{\mathbf{u}_{\mathcal{K}}} \| \mu_{U_{\mathcal{K}}})$$

by assuming that $\mathcal{U}_j \equiv \{0, 1\}$ is a finite field, and A_j is an $l_{\mathcal{A}_j} \times n$ matrix on $\{0, 1\}$ for all $j \in \mathcal{K}$. It is an extension of the operation introduced in Section VI.

Let $t_{\mathbf{u}_{\mathcal{K}}}(b^k) \in \{0, \dots, n\}$ be the number of occurrences of $b^k \in \{0, 1\}^k$ in $\mathbf{u}_{\mathcal{K}} \in \{0, 1\}^{kn}$. We have

$$\hat{g}_{A_{\mathcal{K}}}(\mathbf{a}_{\mathcal{K}}) = \arg \min_{\substack{\mathbf{u}_{\mathcal{K}}, t: \\ t \in \mathcal{H}}} D(\nu_{\mathbf{u}_{\mathcal{K}}, t} \| \mu_{U_{\mathcal{K}}}),$$

where

$$\begin{aligned} \mathbf{u}_{\mathcal{K}, t} &\equiv \begin{cases} \mathbf{u}_{\mathcal{K}} & \text{if } \exists \mathbf{u}_{\mathcal{K}} \text{ s.t. } t_{\mathbf{u}_{\mathcal{K}}} = t, A_j \mathbf{u}_j = \mathbf{a}_j \text{ for all } j \in \mathcal{K} \\ \text{'error'} & \text{otherwise} \end{cases} \\ \mathcal{H} &\equiv \left\{ t : t(b^k) \in \{0, \dots, n\}, \sum_{b^k} t(b^k) = n \right\} \end{aligned} \quad (51)$$

and $D(\nu_{\mathbf{u}_{\mathcal{K}}, t} \| \mu_{U_{\mathcal{K}}}) \equiv \infty$ when the operation (51) returns an error. It should be noted that the cardinality of the set \mathcal{H} is at most $[n+1]^{2^k}$.

We can apply the linear programming technique introduced in [9] to the construction of (51). By using the technique introduced in [9], the conditions $A_j \mathbf{u}_j = \mathbf{a}_j$ for all $j \in \mathcal{K}$ can be replaced by linear inequalities. In the following, we focus on the condition $t_{\mathbf{u}_{\mathcal{K}}} = t$ for a given $t \in \mathcal{H}$.

First, for each $b^k \in \{0, 1\}^k$, we introduce the auxiliary variable $(s_1(b^k), \dots, s_n(b^k))$ defined as

$$s_i(b^k) \equiv \begin{cases} 1, & \text{if } u_{j,i} = b_j \text{ for all } j \in \mathcal{K} \\ 0, & \text{otherwise.} \end{cases} \quad (52)$$

Then the condition $t_{\mathbf{u}_{\mathcal{K}}} = t$ is equivalent to the linear equality

$$\sum_{i=1}^n s_i(b^k) = t(b^k)$$

for all $b^k \in \{0, 1\}^k$.

Next, we show that the relation (52) can be replaced by linear inequalities. Let $\mathcal{S}(b^k) \subset \{0, 1\}^{k+1}$ be defined as

$$\mathcal{S}(b^k) \equiv \{(1, b^k)\} \cup \left[\bigcup_{b'^k \neq b^k} \{(0, b'^k)\} \right]. \quad (53)$$

We have the following lemma, which is proved in Appendix G.

Lemma 9: The convex hull of $\mathcal{S}(b^k)$ is a polytope represented by the following inequalities:

$$v_0 \geq 0$$

$$v_j \geq 0 \text{ for } j \in \mathcal{K} \text{ s.t. } b_j = 0$$

$$v_j \leq 1 \text{ for } j \in \mathcal{K} \text{ s.t. } b_j = 1$$

$$v_0 + [-1]^{b_j} v_j \leq 1 - b_j$$

$$v_0 + \sum_{j \in \mathcal{K}} [-1]^{b_j} v_j \geq 1 - \sum_{j \in \mathcal{K}} b_j$$

where constants b^k and variables v_0, \dots, v_k are considered to be real numbers. Furthermore, there is no non-integral vertex of this polytope.

By using this lemma, we have $[2[k+1]n+1]2^k$ inequalities, which replace the condition $\mathbf{t}_{\mathbf{u}_K} = \mathbf{t}$, defined as

$$\sum_{i=1}^n s_i(b^k) = t(b^k)$$

$$s_i(b^k) \geq 0$$

$$u_{j,i} \geq 0 \text{ for } j \in \mathcal{K} \text{ s.t. } b_j = 0$$

$$u_{j,i} \leq 1 \text{ for } j \in \mathcal{K} \text{ s.t. } b_j = 1$$

$$s_i(b^k) + [-1]^{b_j} u_{j,i} \leq 1 - b_j \text{ for all } j \in \mathcal{K}$$

$$s_i(b^k) + \sum_{j \in \mathcal{K}} [-1]^{b_j} u_{j,i} \geq 1 - \sum_{j \in \mathcal{K}} b_j,$$

where we take all $i \in \{1, 2, \dots, n\}$ and $b^k \in \{0, 1\}^k$.

VIII. PROOF OF THEOREMS

A. Proof of Theorem 1

For a type \mathbf{t} , let $\mathcal{C}_{\mathbf{t}}$ be defined as

$$\mathcal{C}_{\mathbf{t}} \equiv \{\mathbf{u} \in \mathcal{U}^n : \mathbf{t}(\mathbf{u}) = \mathbf{t}\}.$$

We assume that $p_A(\{A : A\mathbf{u} = \mathbf{0}\})$ depends on \mathbf{u} only through the type $\mathbf{t}(\mathbf{u})$. For a given $\mathbf{u} \in \mathcal{C}_{\mathbf{t}}$, we define

$$p_{A,\mathbf{t}} \equiv p_A(\{A : A\mathbf{u} = \mathbf{0}\}).$$

We use the following lemma.

Lemma 10 ([17, Lemma 9]): Let (α_A, β_A) be defined by (9) and (10). Then

$$\alpha_A = |\text{Im} \mathcal{A}| \max_{\mathbf{t} \in \widehat{\mathcal{H}}_A} p_{A,\mathbf{t}} \tag{54}$$

$$\beta_A = \sum_{\mathbf{t} \in \mathcal{H} \setminus \widehat{\mathcal{H}}_A} |\mathcal{C}_{\mathbf{t}}| p_{A,\mathbf{t}}, \tag{55}$$

where \mathcal{H} is a set of all types of length n except the type of the zero vector.

Now we prove Theorem 1. It is enough to show (H3) because (H1), (H2) are satisfied from the assumption of the theorem. Since function A is linear, we have

$$\begin{aligned} p_A(\{A : A\mathbf{u} = A\mathbf{u}'\}) &= p_A(\{A : A[\mathbf{u} - \mathbf{u}'] = \mathbf{0}\}) \\ &= p_{A,\mathbf{t}(\mathbf{u}-\mathbf{u}')} \end{aligned} \tag{56}$$

Then, for $\mathbf{u} \neq \mathbf{u}'$ satisfying $t(\mathbf{u} - \mathbf{u}') \in \widehat{\mathcal{H}}_A$, we have

$$\begin{aligned} p_A(\{A : A\mathbf{u} = A\mathbf{u}'\}) &= p_{A,t(\mathbf{u}-\mathbf{u}')} \\ &\leq \max_{t \in \widehat{\mathcal{H}}_A} p_{A,t} \\ &= \frac{\alpha_A}{|\text{Im}\mathcal{A}|}, \end{aligned} \quad (57)$$

where the last inequality comes from (54). Then we have the fact that $p_A(\{A : A\mathbf{u} = A\mathbf{u}'\}) > \alpha_A/|\text{Im}\mathcal{A}|$ implies $t(\mathbf{u} - \mathbf{u}') \notin \widehat{\mathcal{H}}_A$. Finally, we have

$$\begin{aligned} \sum_{\substack{\mathbf{u}' \in \mathcal{U}^n \setminus \{\mathbf{u}\} \\ p_A(\{A : A\mathbf{u} = A\mathbf{u}'\}) > \frac{\alpha_A}{|\text{Im}\mathcal{A}|}}} p_A(\{A : A\mathbf{u} = A\mathbf{u}'\}) &\leq \sum_{\substack{\mathbf{u}' \in \mathcal{U}^n \setminus \{\mathbf{u}\} \\ t(\mathbf{u}-\mathbf{u}') \in \widehat{\mathcal{H}}_A}} p_{A,t(\mathbf{u}-\mathbf{u}')} \\ &\leq \sum_{t \in \mathcal{H} \setminus \widehat{\mathcal{H}}_A} \sum_{\substack{\mathbf{u}' \in \mathcal{U}^n \setminus \{\mathbf{u}\} \\ t(\mathbf{u}-\mathbf{u}') = t}} p_{A,t} \\ &\leq \sum_{t \in \mathcal{H} \setminus \widehat{\mathcal{H}}_A} |\mathcal{C}_t| p_{A,t} \\ &= \beta_A, \end{aligned} \quad (58)$$

where the equality comes from (55). ■

B. Proof of Theorem 2

Let (\mathbf{x}, \mathbf{y}) be the output of correlated sources. We define

$$\bullet (\mathbf{x}, \mathbf{y}) \notin \mathcal{T}_{XY,\gamma} \quad (\text{SW1})$$

$$\bullet \exists (\mathbf{x}', \mathbf{y}') \neq (\mathbf{x}, \mathbf{y}) \text{ s.t. } \mathbf{x}' \in \mathcal{C}_A(A\mathbf{x}), \mathbf{y}' \in \mathcal{C}_B(B\mathbf{y}), D(\nu_{\mathbf{x}'\mathbf{y}'} \| \mu_{XY}) \leq D(\nu_{\mathbf{x}\mathbf{y}} \| \mu_{XY}). \quad (\text{SW2})$$

It should be noted that (SW2) includes the following three cases:

- there is $(\mathbf{x}', \mathbf{y}')$ satisfying $\mathbf{x}' \neq \mathbf{x}$, $\mathbf{y}' = \mathbf{y}$, $\mathbf{x}' \in \mathcal{C}_A(A\mathbf{x})$, and $D(\nu_{\mathbf{x}'\mathbf{y}'} \| \mu_{XY}) \leq D(\nu_{\mathbf{x}\mathbf{y}} \| \mu_{XY})$,
- there is $(\mathbf{x}', \mathbf{y}')$ satisfying $\mathbf{x}' = \mathbf{x}$, $\mathbf{y}' \neq \mathbf{y}$, $\mathbf{y}' \in \mathcal{C}_B(B\mathbf{y})$, and $D(\nu_{\mathbf{x}'\mathbf{y}'} \| \mu_{XY}) \leq D(\nu_{\mathbf{x}\mathbf{y}} \| \mu_{XY})$,
- there is $(\mathbf{x}', \mathbf{y}')$ satisfying $\mathbf{x}' \neq \mathbf{x}$, $\mathbf{y}' \neq \mathbf{y}$, $\mathbf{x}' \in \mathcal{C}_A(A\mathbf{x})$, $\mathbf{y}' \in \mathcal{C}_B(B\mathbf{y})$, and $D(\nu_{\mathbf{x}'\mathbf{y}'} \| \mu_{XY}) \leq D(\nu_{\mathbf{x}\mathbf{y}} \| \mu_{XY})$.

Since a decoding error occurs when at least one of the conditions (SW1) and (SW2) is satisfied, the error probability is upper bounded by

$$\text{Error}_{XY}(A, B) \leq \mu_{XY}(\mathcal{E}_1) + \mu_{XY}(\mathcal{E}_1^c \cap \mathcal{E}_2), \quad (59)$$

where we define

$$\mathcal{E}_i \equiv \{(\mathbf{x}, \mathbf{y}) : (\text{SW}i)\}.$$

First, we evaluate $E_{AB}[\mu_{XY}(\mathcal{E}_1)]$. From Lemma 22, we have

$$E_{AB}[\mu_{XY}(\mathcal{E}_1)] \leq \frac{\delta}{2} \quad (60)$$

for all sufficiently large n .

Next, we evaluate $E_{AB} [\mu_{XY}(\mathcal{E}_1^c \cap \mathcal{E}_2)]$. When (SW2) is satisfied but (SW1) is not, we have

$$[\mathcal{T}_{XY,\gamma} \setminus \{(\mathbf{x}, \mathbf{y})\}] \cap [\mathcal{C}_A(A\mathbf{x}) \times \mathcal{C}_B(B\mathbf{y})] \neq \emptyset.$$

Applying Lemma 5 by letting $\mathcal{G} \equiv \mathcal{T}_{XY,\gamma}$, we have

$$\begin{aligned} & E_{AB} [\mu_{XY}(\mathcal{E}_1^c \cap \mathcal{E}_2)] \\ &= \sum_{(\mathbf{x}, \mathbf{y}) \in \mathcal{T}_{XY,\gamma}} \mu_{XY}(\mathbf{x}, \mathbf{y}) p_{AB}(\{(A, B) : (\text{SW2})\}) \\ &\leq \sum_{(\mathbf{x}, \mathbf{y}) \in \mathcal{T}_{XY,\gamma}} \mu_{XY}(\mathbf{x}, \mathbf{y}) p_{AB} \left(\left\{ (A, B) : \right. \right. \\ &\quad \left. \left. [\mathcal{T}_{XY,\gamma} \setminus \{(\mathbf{x}, \mathbf{y})\}] \cap [\mathcal{C}_A(A\mathbf{x}) \times \mathcal{C}_B(B\mathbf{y})] \neq \emptyset \right\} \right) \\ &\leq \frac{|\mathcal{G}| \alpha_A \alpha_B}{|\text{Im}\mathcal{A}| |\text{Im}\mathcal{B}|} + \frac{\left[\max_{\mathbf{y} \in \mathcal{G}_Y} |\mathcal{G}_{X|Y}(\mathbf{y})| \right] \alpha_A [\beta_B + 1]}{|\text{Im}\mathcal{A}|} + \frac{\left[\max_{\mathbf{x} \in \mathcal{G}_X} |\mathcal{G}_{Y|X}(\mathbf{x})| \right] \alpha_B [\beta_A + 1]}{|\text{Im}\mathcal{B}|} + \beta_A + \beta_B + \beta_A \beta_B \\ &\leq \frac{|\mathcal{T}_{XY,\gamma}| \alpha_A \alpha_B}{|\text{Im}\mathcal{A}| |\text{Im}\mathcal{B}|} + \frac{\left[\max_{\mathbf{y} \in \mathcal{T}_{Y,\gamma}} |\mathcal{T}_{X|Y,\gamma}(\mathbf{y})| \right] \alpha_A [\beta_B + 1]}{|\text{Im}\mathcal{A}|} + \frac{\left[\max_{\mathbf{x} \in \mathcal{T}_{X,\gamma}} |\mathcal{T}_{Y|X,\gamma}(\mathbf{x})| \right] \alpha_B [\beta_A + 1]}{|\text{Im}\mathcal{B}|} + \beta_A + \beta_B + \beta_A \beta_B \\ &\leq \frac{2^{n[H(X,Y) + \eta_{XY}(\gamma)]} \alpha_A \alpha_B}{|\text{Im}\mathcal{A}| |\text{Im}\mathcal{B}|} + \frac{2^{n[H(X|Y) + \eta_{X|Y}(\gamma|\gamma)]} \alpha_A [\beta_B + 1]}{|\text{Im}\mathcal{A}|} + \frac{2^{n[H(Y|X) + \eta_{Y|X}(\gamma|\gamma)]} \alpha_B [\beta_A + 1]}{|\text{Im}\mathcal{B}|} \\ &\quad + \beta_A + \beta_B + \beta_A \beta_B \\ &= 2^{-n[R_X + R_Y - H(X,Y) - \eta_{XY}(\gamma)]} \alpha_A \alpha_B + 2^{-n[R_X - H(X|Y) - \eta_{X|Y}(\gamma|\gamma)]} \alpha_A [\beta_B + 1] \\ &\quad + 2^{-n[R_Y - H(Y|X) - \eta_{Y|X}(\gamma|\gamma)]} \alpha_B [\beta_A + 1] + \beta_A + \beta_B + \beta_A \beta_B \\ &\leq \frac{\delta}{4} \end{aligned} \tag{61}$$

for all sufficiently large n by taking an appropriate $\gamma > 0$, where the third inequality comes from the fact that $\mathcal{G}_X \subset \mathcal{T}_{X,\gamma}$, $\mathcal{G}_Y \subset \mathcal{T}_{Y,\gamma}$, $\mathcal{G}_{X|Y}(\mathbf{y}) \subset \mathcal{T}_{X|Y,\gamma}(\mathbf{y})$, and $\mathcal{G}_{Y|X}(\mathbf{x}) \subset \mathcal{T}_{Y|X,\gamma}(\mathbf{x})$, which are obtained from Lemma 20. The fourth inequality comes from Lemma 23 and the last inequality comes from (11)–(13) and conditions (H1), (H2) of ensembles (\mathcal{A}, p_A) and (\mathcal{B}, p_B) .

Finally, from (59)–(61), for all $\delta > 0$ and for all sufficiently large n there are A and B such that

$$\text{Error}_{XY}(A, B) < \delta.$$

■

Remark 7: We need the conditions $\mathbf{x}' \in \mathcal{T}_{X,\gamma}$ and $\mathbf{y}' \in \mathcal{T}_{Y,\gamma}$ in the maximum-likelihood decoder (14) because $\mathcal{G} \equiv \{(\mathbf{x}', \mathbf{y}') : \mu_{XY}(\mathbf{x}', \mathbf{y}') \geq \mu_{XY}(\mathbf{x}, \mathbf{y})\}$ does not satisfy

$$\begin{aligned} \max_{\mathbf{y} \in \mathcal{G}_Y} |\mathcal{G}_{X|Y}(\mathbf{y})| &\leq 2^{n[H(X|Y) + \eta_{X|Y}(\gamma|\gamma)]} \\ \max_{\mathbf{x} \in \mathcal{G}_X} |\mathcal{G}_{Y|X}(\mathbf{x})| &\leq 2^{n[H(Y|X) + \eta_{Y|X}(\gamma|\gamma)]} \end{aligned}$$

in general. It is unknown if we can remove these conditions. It seems that the same problem appears in [16, Eq. (20)].

C. Proof of Theorem 3

In the following, we assume that ensembles $(\mathcal{A}, p_{\mathcal{A}})$, $(\mathcal{A}', p_{\mathcal{A}'})$, $(\mathcal{B}, p_{\mathcal{B}})$, and $(\mathcal{B}', p_{\mathcal{B}'})$ have a strong hash property. Then, from Lemma 4, ensembles $(\hat{\mathcal{A}}, p_{\hat{\mathcal{A}}})$ and $(\hat{\mathcal{B}}, p_{\hat{\mathcal{B}}})$, defined by

$$\hat{A}u \equiv (Au, A'u)$$

$$\hat{B}v \equiv (Bv, B'v)$$

have an $(\alpha_{\hat{\mathcal{A}}}, \beta_{\hat{\mathcal{A}}})$ -hash property and an $(\alpha_{\hat{\mathcal{B}}}, \beta_{\hat{\mathcal{B}}})$ -hash property, respectively, where

$$|\text{Im}\hat{\mathcal{A}}| = |\text{Im}\mathcal{A}||\text{Im}\mathcal{A}'|$$

$$|\text{Im}\hat{\mathcal{B}}| = |\text{Im}\mathcal{B}||\text{Im}\mathcal{B}'|$$

$$\alpha_{\hat{\mathcal{A}}} \equiv \alpha_{\mathcal{A}}\alpha_{\mathcal{A}'}$$

$$\beta_{\hat{\mathcal{A}}} \equiv \beta_{\mathcal{A}} + \beta_{\mathcal{A}'}$$

$$\alpha_{\hat{\mathcal{B}}} \equiv \alpha_{\mathcal{B}}\alpha_{\mathcal{B}'}$$

$$\beta_{\hat{\mathcal{B}}} \equiv \beta_{\mathcal{B}} + \beta_{\mathcal{B}'}$$

For $\beta_{\mathcal{A}}$ and $\beta_{\mathcal{B}}$ satisfying

$$\lim_{n \rightarrow \infty} \beta_{\mathcal{A}}(n) = 0$$

$$\lim_{n \rightarrow \infty} \beta_{\mathcal{B}}(n) = 0,$$

let $\kappa \equiv \{\kappa(n)\}_{n=1}^{\infty}$ be a sequence satisfying

$$\lim_{n \rightarrow \infty} \kappa(n) = \infty \tag{62}$$

$$\lim_{n \rightarrow \infty} \kappa(n)\beta_{\mathcal{A}}(n) = 0 \tag{63}$$

$$\lim_{n \rightarrow \infty} \kappa(n)\beta_{\mathcal{B}}(n) = 0 \tag{64}$$

$$\lim_{n \rightarrow \infty} \frac{\log \kappa(n)}{n} = 0. \tag{65}$$

For example, we obtain such a κ by letting

$$\kappa(n) \equiv \begin{cases} n^{\xi} & \text{if } \beta(n) = o(n^{-\xi}), \xi > 0 \\ \frac{1}{\sqrt{\beta(n)}}, & \text{otherwise} \end{cases}$$

for every n , where $\beta(n) \equiv \max\{\beta_{\mathcal{A}}(n), \beta_{\mathcal{B}}(n)\}$. If $\beta(n)$ is not $o(n^{-\xi})$, there is $\kappa' > 0$ such that $\beta(n)n^{\xi} > \kappa'$ and

$$\begin{aligned} \frac{\log \kappa(n)}{n} &= \frac{\log \frac{1}{\sqrt{\beta(n)}}}{2n} \\ &\leq \frac{\log \frac{n^{\xi}}{\kappa'}}{2n} \\ &= \frac{\xi \log n - \log \kappa'}{2n} \end{aligned} \tag{66}$$

for all sufficiently large n . This implies that κ satisfies (65). In the following, κ denotes $\kappa(n)$.

From (20)–(25), and (65), we have the fact that there is $\gamma > 0$ such that

$$\varepsilon \geq \eta_{\mathcal{U}\mathcal{V}}(\gamma) + \frac{\log \kappa}{n} \quad (67)$$

$$r_Y > H(U|Y) + \zeta_{\mathcal{U}|Y}(2\gamma|2\gamma) \quad (68)$$

$$r_Z > H(V|Z) + \zeta_{\mathcal{V}|Z}(2\gamma|2\gamma) \quad (69)$$

$$r_Y + R_Y \leq H(U) - \varepsilon - \eta_{\mathcal{V}|\mathcal{U}}(\gamma|\gamma) \quad (70)$$

$$r_Z + R_Z \leq H(V) - \varepsilon - \eta_{\mathcal{U}|\mathcal{V}}(\gamma|\gamma) \quad (71)$$

$$r_Y + R_Y + r_Z + R_Z \leq H(U, V) - \eta_{\mathcal{U}\mathcal{V}}(\gamma) - \frac{\log \kappa}{n} \quad (72)$$

$$r_Y + R_Y + r_Z + R_Z \geq H(U, V) - \varepsilon \quad (73)$$

for all sufficiently large n .

We have

$$\begin{aligned} |\mathcal{T}_{UV,\gamma}| &\geq 2^{n[H(U,V) - \eta_{\mathcal{U}\mathcal{V}}(\gamma)]} \\ &\geq \kappa 2^{n[r_Y + R_Y + r_Z + R_Z]} \\ &= \kappa |\text{Im}\mathcal{A}| |\text{Im}\mathcal{A}'| |\text{Im}\mathcal{B}| |\text{Im}\mathcal{B}'| \\ &= \kappa |\text{Im}\hat{\mathcal{A}}| |\text{Im}\hat{\mathcal{B}}| \end{aligned} \quad (74)$$

for all sufficiently large n , where the first inequality comes from Lemma 23, and the second inequality comes from (72). This implies that there is $\mathcal{T} \subset \mathcal{T}_{UV,\gamma}$ such that

$$\kappa \leq \frac{|\mathcal{T}|}{|\text{Im}\hat{\mathcal{A}}| |\text{Im}\hat{\mathcal{B}}|} \leq 2\kappa \quad (75)$$

for all sufficiently large n , where we construct such \mathcal{T} by taking $|\mathcal{T}|$ elements from $\mathcal{T}_{UV,\gamma}$ in descending order of divergence.

Now we prove the theorem. Let \mathbf{m} and \mathbf{w} be messages. Let \mathbf{u} and \mathbf{v} denotes the \mathcal{U}^n -component (resp. \mathcal{V}^n -component) of $g_{AA'BB'}(\mathbf{a}, \mathbf{m}, \mathbf{b}, \mathbf{w})$, that is,

$$(\mathbf{u}, \mathbf{v}) \equiv \hat{g}_{AA'BB'}(\mathbf{a}, \mathbf{m}, \mathbf{b}, \mathbf{w}).$$

Let \mathbf{y} and \mathbf{z} be the channel outputs. We define

$$\bullet (\mathbf{u}, \mathbf{v}) \in \mathcal{T} \subset \mathcal{T}_{UV,\gamma} \quad (\text{BC1})$$

$$\bullet (\mathbf{y}, \mathbf{z}) \in \mathcal{T}_{YZ|UVX,\gamma}(\mathbf{u}, \mathbf{v}, \mathbf{x}) \quad (\text{BC2})$$

$$\bullet g_A(\mathbf{a}|\mathbf{y}) = \mathbf{u} \quad (\text{BC3})$$

$$\bullet g_B(\mathbf{b}|\mathbf{z}) = \mathbf{v}, \quad (\text{BC4})$$

where we define $\mathbf{x} \equiv \mathbf{f}(\mathbf{u}, \mathbf{v})$. Then the error probability (34) is upper bounded by

$$\begin{aligned} &\text{Error}_{YZ|X}(A, A', B, B', \mathbf{a}, \mathbf{b}) \\ &\leq p_{\mathbf{m}\mathbf{w}YZ}(\mathcal{S}_1^c) + p_{\mathbf{m}\mathbf{w}YZ}(\mathcal{S}_2^c) + p_{\mathbf{m}\mathbf{w}YZ}(\mathcal{S}_1 \cap \mathcal{S}_2 \cap \mathcal{S}_3^c) + p_{\mathbf{m}\mathbf{w}YZ}(\mathcal{S}_1 \cap \mathcal{S}_2 \cap \mathcal{S}_4^c), \end{aligned} \quad (76)$$

where

$$\mathcal{S}_i \equiv \{(\mathbf{m}, \mathbf{w}, \mathbf{y}, \mathbf{z}) : (\text{BC}i)\}.$$

In the following, we define

$$\hat{\mathbf{a}} \equiv (\mathbf{a}, \mathbf{m})$$

$$\hat{\mathbf{b}} \equiv (\mathbf{b}, \mathbf{w})$$

and

$$\mathcal{C}_{\hat{A}\hat{B}}(\hat{\mathbf{a}}, \hat{\mathbf{b}}) \equiv [\mathcal{C}_A(\mathbf{a}) \cap \mathcal{C}_{A'}(\mathbf{m})] \times [\mathcal{C}_B(\mathbf{b}) \cap \mathcal{C}_{B'}(\mathbf{w})].$$

First, we evaluate $E_{AA'BB'\mathbf{ab}}[p_{\mathbf{mwYZ}}(\mathcal{S}_1^c)]$. From Lemma 6, we have

$$\begin{aligned} & E_{AA'BB'\mathbf{ab}}[p_{\mathbf{mwYZ}}(\mathcal{S}_1^c)] \\ &= p_{AA'BB'\mathbf{abmw}}(\{(A, A', B, B', \mathbf{a}, \mathbf{b}, \mathbf{m}, \mathbf{w}) : \hat{g}_{AA'BB'}(\mathbf{a}, \mathbf{m}, \mathbf{b}, \mathbf{w}) \notin \mathcal{T}\}) \\ &\leq p_{\hat{A}\hat{B}\hat{\mathbf{a}}\hat{\mathbf{b}}}(\{(\hat{A}, \hat{B}, \hat{\mathbf{a}}, \hat{\mathbf{b}}) : \mathcal{T} \cap \mathcal{C}_{\hat{A}\hat{B}}(\hat{\mathbf{a}}, \hat{\mathbf{b}}) = \emptyset\}) \\ &\leq \alpha_{\hat{A}}\alpha_{\hat{B}} - 1 + \frac{|\text{Im}\hat{\mathcal{B}}| \left[\max_{\mathbf{v} \in \mathcal{T}_V} |\mathcal{T}_{\mathcal{U}|\mathcal{V}}(\mathbf{v})| \right] \alpha_{\hat{A}}[\beta_{\hat{B}} + 1]}{|\mathcal{T}|} + \frac{|\text{Im}\hat{\mathcal{A}}| \left[\max_{\mathbf{u} \in \mathcal{T}_U} |\mathcal{T}_{\mathcal{V}|\mathcal{U}}(\mathbf{u})| \right] \alpha_{\hat{B}}[\beta_{\hat{A}} + 1]}{|\mathcal{T}|n} \\ &\quad + \frac{|\text{Im}\hat{\mathcal{A}}||\text{Im}\hat{\mathcal{B}}| [\beta_{\hat{A}} + \beta_{\hat{B}} + \beta_{\hat{A}}\beta_{\hat{B}} + 1]}{|\mathcal{T}|}. \end{aligned} \quad (77)$$

From Lemma 20 and the fact that $\mathcal{T} \subset \mathcal{T}_{XY,\gamma}$, we have the fact that $\mathbf{u} \in \mathcal{T}_U$ implies $\mathbf{u} \in \mathcal{T}_{U,\gamma}$ and $\mathbf{v} \in \mathcal{T}_{\mathcal{V}|\mathcal{U}}(\mathbf{u})$ implies $\mathbf{v} \in \mathcal{T}_{\mathcal{V}|U,\gamma}(\mathbf{u})$. Furthermore, from Lemma 23, we have

$$\begin{aligned} \max_{\mathbf{v} \in \mathcal{T}_V} |\mathcal{T}_{\mathcal{U}|\mathcal{V}}(\mathbf{v})| &\leq \max_{\mathbf{v} \in \mathcal{T}_{V,\gamma}} |\mathcal{T}_{U|V,\gamma}(\mathbf{v})| \\ &\leq 2^{n[H(U|V) + \eta_{U|V}(\gamma|\gamma)]}. \end{aligned} \quad (78)$$

Then we have

$$\begin{aligned} \frac{|\text{Im}\hat{\mathcal{B}}| \max_{\mathbf{v} \in \mathcal{T}_V} |\mathcal{T}_{\mathcal{U}|\mathcal{V}}(\mathbf{v})|}{|\mathcal{T}|} &\leq \frac{2^{n[H(U|V) + \eta_{U|V}(\gamma|\gamma)]}}{\kappa |\text{Im}\hat{\mathcal{A}}|} \\ &= \frac{2^{-n[r_Y + R_Y - H(U|V) - \eta_{U|V}(\gamma|\gamma)]}}{\kappa} \\ &\leq \frac{1}{\kappa} \end{aligned} \quad (79)$$

where the first inequality comes from (75), (78), and the second inequality comes from the fact that

$$\begin{aligned} r_Y + R_Y &= r_Y + R_Y + r_Z + R_Z - r_Z - R_Z \\ &\geq H(U, V) - \varepsilon - r_Z - R_Z \\ &\geq H(U, V) - \varepsilon - H(V) + \varepsilon + \eta_{U|V}(\gamma|\gamma) \\ &= H(U|V) + \eta_{U|V}(\gamma|\gamma) \end{aligned} \quad (80)$$

which is obtained from (70) and (73). Similarly, we have

$$\frac{|\text{Im}\hat{\mathcal{A}}| \max_{\mathbf{u} \in \mathcal{T}_U} |\mathcal{T}_{\mathcal{V}|\mathcal{U}}(\mathbf{u})|}{|\mathcal{T}|} \leq \frac{2^{-n[r_Z + R_Z - H(V|U) - \eta_{V|U}(\gamma|\gamma)]}}{\kappa}$$

$$\leq \frac{1}{\kappa} \quad (81)$$

Furthermore, we have

$$\frac{|\text{Im}\hat{\mathcal{A}}||\text{Im}\hat{\mathcal{B}}|}{|\mathcal{T}|} \leq \frac{1}{\kappa}. \quad (82)$$

from (75). Then, from (62), (77), (79) (81), (82), and the properties (H1) and (H2) for ensembles $(\hat{\mathcal{A}}, p_{\hat{\mathcal{A}}})$ and $(\hat{\mathcal{B}}, p_{\hat{\mathcal{B}}})$, we have

$$E_{AA'BB'ab} [p_{\mathbf{m}WYZ}(\mathcal{S}_1^c)] \leq \frac{\delta}{4} \quad (83)$$

for all $\delta > 0$ and sufficiently large n .

Next, we evaluate $E_{AA'BB'ab} [p_{\mathbf{m}WYZ}(\mathcal{S}_2^c)]$. We have

$$\begin{aligned} E_{AA'BB'ab} [p_{\mathbf{m}WYZ}(\mathcal{S}_2^c)] &= E_{\hat{\mathcal{A}}\hat{\mathcal{B}}\hat{\mathcal{A}}\hat{\mathcal{B}}} [\mu_{YZ|X} ([\mathcal{T}_{YZ|UVX,\gamma}(U, V, X)]^c | X)] \\ &= E_{\hat{\mathcal{A}}\hat{\mathcal{B}}\hat{\mathcal{A}}\hat{\mathcal{B}}} [\mu_{YZ|UVX} ([\mathcal{T}_{YZ|UVX,\gamma}(U, V, X)]^c | U, V, X)] \\ &\leq 2^{-n[\gamma - \lambda_{UVX,\gamma}]} \\ &\leq \frac{\delta}{4} \end{aligned} \quad (84)$$

for all $\delta > 0$ and sufficiently large n , where we define $\mathbf{X} \equiv \mathbf{f}(U, V)$, the second equality comes from (18), and the first inequality comes from Lemma 22.

Next, we evaluate $E_{AA'BB'ab} [p_{\mathbf{m}WYZ}(\mathcal{S}_1 \cap \mathcal{S}_2 \cap \mathcal{S}_3^c)]$ and $E_{AA'BB'ab} [p_{\mathbf{m}WYZ}(\mathcal{S}_1 \cap \mathcal{S}_2 \cap \mathcal{S}_4^c)]$. In the following, we assume (BC1), (BC2), and $g_A(\mathbf{a}|\mathbf{y}) \neq \mathbf{u}$, where the last assumption is equivalent to $g_A(A\mathbf{u}|\mathbf{y}) \neq \mathbf{u}$ from (BC1). From (19) and Lemma 20, we have the fact that $(\mathbf{u}, \mathbf{v}) \in \mathcal{T}$ implies $(\mathbf{u}, \mathbf{v}, \mathbf{x}) \in \mathcal{T}_{UVX,\gamma}$. From Lemma 20, we have $(\mathbf{u}, \mathbf{y}) \in \mathcal{T}_{UY,2\gamma}$ and $\mathbf{u} \in \mathcal{T}_{U|Y,2\gamma}(\mathbf{y})$. Then there is $\mathbf{u}' \in \mathcal{C}_A(A\mathbf{u})$ such that $\mathbf{u}' \neq \mathbf{u}$ and

$$\begin{aligned} \mu_{U|Y}(\mathbf{u}'|\mathbf{y}) &\geq \mu_{U|Y}(\mathbf{u}|\mathbf{y}) \\ &\geq 2^{-n[H(U|Y) + \zeta_{U|Y}(2\gamma|2\gamma)]}, \end{aligned} \quad (85)$$

where the second inequality comes from Lemma 22. This implies that $[\mathcal{G}(\mathbf{y}) \setminus \{\mathbf{u}\}] \cap \mathcal{C}_A(A\mathbf{u}) \neq \emptyset$, where

$$\mathcal{G}(\mathbf{y}) \equiv \left\{ \mathbf{u}' : \mu_{U|Y}(\mathbf{u}'|\mathbf{y}) \geq 2^{-[H(U|Y) + \zeta_{U|Y}(2\gamma|2\gamma)]} \right\}.$$

From Lemma 2, we have

$$\begin{aligned} E_A [\chi(g_A(A\mathbf{u}|\mathbf{y}) \neq \mathbf{u})] &\leq p_A(\{A : [\mathcal{G}(\mathbf{y}) \setminus \{\mathbf{u}\}] \cap \mathcal{C}_A(A\mathbf{u}) \neq \emptyset\}) \\ &\leq \frac{|\mathcal{G}(\mathbf{y})|_{\alpha_A}}{|\text{Im}\mathcal{A}|} + \beta_A \\ &\leq 2^{-n[r_Y - H(U|Y) - \zeta_{U|Y}(2\gamma|2\gamma)]} \alpha_A + \beta_A. \end{aligned} \quad (86)$$

the last inequality comes from the definition of r_Y and the fact that

$$|\mathcal{G}(\mathbf{y})| \leq 2^{n[H(U|Y) + \zeta_{U|Y}(2\gamma|2\gamma)]}.$$

Then we have

$$\begin{aligned} &E_{AA'BB'ab} [p_{\mathbf{m}WYZ}(\mathcal{S}_1 \cap \mathcal{S}_2 \cap \mathcal{S}_3^c)] \\ &\leq E_{AA'BB'ab\mathbf{m}W} \left[\sum_{(\mathbf{u}, \mathbf{v}) \in \mathcal{T}} \chi(\hat{g}_{AA'BB'}(\mathbf{a}, \mathbf{m}, \mathbf{b}, \mathbf{w}) = (\mathbf{u}, \mathbf{v})) \sum_{\mathbf{y} \in \mathcal{T}_{Y|UVX,\gamma}(\mathbf{u}, \mathbf{v}, \mathbf{x})} \mu_{Y|X}(\mathbf{y}|\mathbf{x}) \chi(g_A(\mathbf{a}|\mathbf{y}) \neq \mathbf{u}) \right] \end{aligned}$$

$$\begin{aligned}
&\leq E_{AA'BB'abmw} \left[\sum_{(\mathbf{u}, \mathbf{v}) \in \mathcal{T}} \chi(A\mathbf{u} = \mathbf{a}) \chi(A'\mathbf{u} = \mathbf{m}) \chi(B\mathbf{v} = \mathbf{b}) \chi(B'\mathbf{v} = \mathbf{w}) \right. \\
&\quad \cdot \left. \sum_{\mathbf{y} \in \mathcal{T}_{Y|UVX, \gamma}(\mathbf{u}, \mathbf{v}, \mathbf{x})} \mu_{Y|X}(\mathbf{y}|\mathbf{x}) \chi(g_A(\mathbf{a}|\mathbf{y}) \neq \mathbf{u}) \right] \\
&= \sum_{(\mathbf{u}, \mathbf{v}) \in \mathcal{T}} \sum_{\mathbf{y} \in \mathcal{T}_{Y|UVX, \gamma}(\mathbf{u}, \mathbf{v}, \mathbf{x})} \mu_{Y|X}(\mathbf{y}|\mathbf{x}) \\
&\quad \cdot E_A [\chi(g_A(A\mathbf{u}|\mathbf{y}) \neq \mathbf{u}) E_{A'BB'abmw} [\chi(A\mathbf{u} = \mathbf{a}) \chi(A'\mathbf{u} = \mathbf{m}) \chi(B\mathbf{v} = \mathbf{b}) \chi(B'\mathbf{v} = \mathbf{w})]]] \\
&= \frac{1}{|\text{Im}\mathcal{A}||\text{Im}\mathcal{A}'||\text{Im}\mathcal{B}||\text{Im}\mathcal{B}'|} \sum_{(\mathbf{u}, \mathbf{v}) \in \mathcal{T}} \sum_{\mathbf{y} \in \mathcal{T}_{Y|UVX, \gamma}(\mathbf{u}, \mathbf{v}, \mathbf{x})} \mu_{Y|X}(\mathbf{y}|\mathbf{x}) E_A [\chi(g_A(A\mathbf{u}|\mathbf{y}) \neq \mathbf{u})] \\
&\leq \left[2^{-n[r_Y - H(U|Y) - \zeta_{U|Y}(2\gamma|2\gamma)]} \alpha_A + \beta_A \right] \sum_{(\mathbf{u}, \mathbf{v}) \in \mathcal{T}} \frac{1}{|\text{Im}\mathcal{A}||\text{Im}\mathcal{A}'||\text{Im}\mathcal{B}||\text{Im}\mathcal{B}'|} \\
&\leq 2\kappa \left[2^{-n[r_Y - H(U|Y) - \zeta_{U|Y}(2\gamma|2\gamma)]} \alpha_A + \beta_A \right] \\
&\leq \frac{\delta}{4} \tag{87}
\end{aligned}$$

for all $\delta > 0$ and sufficiently large n , where the second equality comes from Lemma 11 that appears in Appendix A, the fourth inequality comes from (75), and the last inequality comes from (63), (68), and the conditions (H1), (H2) of (\mathcal{A}, p_A) . Similarly, we have

$$\begin{aligned}
E_{AA'BB'ab} [p_{\mathbf{mw}YZ}(\mathcal{S}_1 \cap \mathcal{S}_2 \cap \mathcal{S}_4^c)] &\leq 2\kappa \left[2^{-n[r_Z - H(V|Z) - \zeta_{V|Z}(2\gamma|2\gamma)]} \alpha_B + \beta_B \right] \\
&\leq \frac{\delta}{4} \tag{88}
\end{aligned}$$

for all $\delta > 0$ and sufficiently large n .

Finally, from (76), (83), (84), (87), and (88), we have the fact that for all $\delta > 0$ and sufficiently large n there are $A \in \mathcal{A}$, $A' \in \mathcal{A}'$, $B \in \mathcal{B}$, $B' \in \mathcal{B}'$, $\mathbf{a} \in \text{Im}\mathcal{A}$, and $\mathbf{b} \in \text{Im}\mathcal{B}$ satisfying (34). ■

IX. CONCLUSION

The constructions of the Slepian-Wolf source code and the broadcast channel code were presented. The proof of the theorems is based on the notion of a strong hash property for an ensemble of functions, where two lemmas called ‘collision-resistance property’ and ‘saturation property’ introduced [17] are extended from a single domain to multiple domains. Since an ensemble of sparse matrices has a strong hash property, we can construct codes by using sparse matrices and it is expected that we can use the efficient approximation algorithms for encoding/decoding. It should be noted that the capacity region for the general broadcast channel coding is unknown and we hope that our approach give us a hint for deriving the general capacity region.

APPENDIX

A. Basic Property of Ensemble

We review the following lemma, which is proved in [17].

Lemma 11 ([17, Lemma 9]): Assume that random variables A and \mathbf{a} are independent. Then,

$$E_{\mathbf{a}} [\chi(A\mathbf{u} = \mathbf{a})] = \frac{1}{|\text{Im}\mathcal{A}|}$$

for any A and \mathbf{u} and

$$E_{\mathbf{A}\mathbf{a}}[\chi(\mathbf{A}\mathbf{u} = \mathbf{a})] = \frac{1}{|\text{Im}\mathcal{A}|}$$

for any \mathbf{u} .

B. Proof of Lemma 1

If an ensemble satisfies (H4), then we have

$$\begin{aligned} & \sum_{\substack{\mathbf{u} \in \mathcal{T} \\ \mathbf{u}' \in \mathcal{T}'}} p_{\mathbf{A}}(\{A : A\mathbf{u} = A\mathbf{u}'\}) \\ &= \sum_{\mathbf{u} \in \mathcal{T} \cap \mathcal{T}'} p_{\mathbf{A}}(\{A : A\mathbf{u} = A\mathbf{u}'\}) + \sum_{\mathbf{u} \in \mathcal{T}} \sum_{\substack{\mathbf{u}' \in \mathcal{T}' \setminus \{\mathbf{u}\} \\ p_{\mathbf{A}}(\{A : A\mathbf{u} = A\mathbf{u}'\}) \leq \frac{\alpha_{\mathbf{A}}}{|\text{Im}\mathcal{A}|}}} p_{\mathbf{A}}(\{A : A\mathbf{u} = A\mathbf{u}'\}) \\ &+ \sum_{\mathbf{u} \in \mathcal{T}} \sum_{\substack{\mathbf{u}' \in \mathcal{T}' \setminus \{\mathbf{u}\} \\ p_{\mathbf{A}}(\{A : A\mathbf{u} = A\mathbf{u}'\}) > \frac{\alpha_{\mathbf{A}}}{|\text{Im}\mathcal{A}|}}} p_{\mathbf{A}}(\{A : A\mathbf{u} = A\mathbf{u}'\}) \\ &\leq |\mathcal{T} \cap \mathcal{T}'| + \sum_{\mathbf{u} \in \mathcal{T}} \sum_{\substack{\mathbf{u}' \in \mathcal{T}' \setminus \{\mathbf{u}\} \\ p_{\mathbf{A}}(\{A : A\mathbf{u} = A\mathbf{u}'\}) \leq \frac{\alpha_{\mathbf{A}}}{|\text{Im}\mathcal{A}|}}} \frac{\alpha_{\mathbf{A}}}{|\text{Im}\mathcal{A}|} + \sum_{\mathbf{u} \in \mathcal{T}} \beta_{\mathbf{A}} \\ &\leq |\mathcal{T} \cap \mathcal{T}'| + \frac{|\mathcal{T}||\mathcal{T}'|\alpha_{\mathbf{A}}}{|\text{Im}\mathcal{A}|} + |\mathcal{T}|\beta_{\mathbf{A}} \\ &\leq |\mathcal{T} \cap \mathcal{T}'| + \frac{|\mathcal{T}||\mathcal{T}'|\alpha_{\mathbf{A}}}{|\text{Im}\mathcal{A}|} + \min\{|\mathcal{T}|, |\mathcal{T}'|\}\beta_{\mathbf{A}} \end{aligned} \tag{89}$$

for any \mathcal{T} and \mathcal{T}' satisfying $|\mathcal{T}| \leq |\mathcal{T}'|$. ■

C. Proof of Lemma 4

Let

$$\begin{aligned} p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} &\equiv p_{\mathbf{A}}(\{A : A\mathbf{u} = A\mathbf{u}'\}) \\ p_{\mathbf{A}', \mathbf{u}, \mathbf{u}'} &\equiv p_{\mathbf{A}'}(\{A' : A'\mathbf{u} = A'\mathbf{u}'\}). \end{aligned}$$

Then we have

$$\begin{aligned} & \sum_{\substack{\mathbf{u}' \in \mathcal{U}^n \setminus \{\mathbf{u}\} \\ p_{\hat{\mathbf{A}}, \mathbf{u}, \mathbf{u}'} > \frac{\alpha_{\hat{\mathbf{A}}}}{|\text{Im}\hat{\mathcal{A}}|}}} p_{\hat{\mathbf{A}}}(\{\hat{A} : \hat{A}\mathbf{u} = \hat{A}\mathbf{u}'\}) \\ &= \sum_{\substack{\mathbf{u}' \in \mathcal{U}^n \setminus \{\mathbf{u}\} \\ p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} p_{\mathbf{B}, \mathbf{u}, \mathbf{u}'} > \frac{\alpha_{\mathbf{A}} \alpha_{\mathbf{B}'}}{|\text{Im}\mathcal{A}| |\text{Im}\mathcal{B}|}}} p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} p_{\mathbf{A}', \mathbf{u}, \mathbf{u}'} \\ &= \sum_{\substack{\mathbf{u}' \in \mathcal{U}^n \setminus \{\mathbf{u}\} \\ p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} p_{\mathbf{A}', \mathbf{u}, \mathbf{u}'} > \frac{\alpha_{\mathbf{A}} \alpha_{\mathbf{A}'}}{|\text{Im}\mathcal{A}| |\text{Im}\mathcal{A}'|} \\ p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} > \frac{\alpha_{\mathbf{A}}}{|\text{Im}\mathcal{A}|}}} p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} p_{\mathbf{A}', \mathbf{u}, \mathbf{u}'} + \sum_{\substack{\mathbf{u}' \in \mathcal{U}^n \setminus \{\mathbf{u}\} \\ p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} p_{\mathbf{A}', \mathbf{u}, \mathbf{u}'} > \frac{\alpha_{\mathbf{A}} \alpha_{\mathbf{A}'}}{|\text{Im}\mathcal{A}| |\text{Im}\mathcal{A}'|} \\ p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} \leq \frac{\alpha_{\mathbf{A}}}{|\text{Im}\mathcal{A}|}}} p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} p_{\mathbf{A}', \mathbf{u}, \mathbf{u}'} \\ &\leq \sum_{\substack{\mathbf{u}' \in \mathcal{U}^n \setminus \{\mathbf{u}\} \\ p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} > \frac{\alpha_{\mathbf{A}}}{|\text{Im}\mathcal{A}|}}} p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} p_{\mathbf{A}', \mathbf{u}, \mathbf{u}'} + \sum_{\substack{\mathbf{u}' \in \mathcal{U}^n \setminus \{\mathbf{u}\} \\ p_{\mathbf{A}', \mathbf{u}, \mathbf{u}'} > \frac{\alpha_{\mathbf{A}'}}{|\text{Im}\mathcal{A}'|}}} p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} p_{\mathbf{A}', \mathbf{u}, \mathbf{u}'} \end{aligned}$$

$$\begin{aligned}
&\leq \sum_{\substack{\mathbf{u}' \in \mathcal{U}^n \setminus \{\mathbf{u}\} \\ p_{A,\mathbf{u},\mathbf{u}'} > \frac{\alpha_A}{|\text{Im}\mathcal{A}|}}} p_{A,\mathbf{u},\mathbf{u}'} + \sum_{\substack{\mathbf{u}' \in \mathcal{U}^n \setminus \{\mathbf{u}\} \\ p_{A',\mathbf{u},\mathbf{u}'} > \frac{\alpha_{A'}}{|\text{Im}\mathcal{A}'|}}} p_{A',\mathbf{u},\mathbf{u}'} \\
&= \beta_A + \beta_{A'} \\
&= \beta_{\hat{A}},
\end{aligned} \tag{90}$$

where the first equality comes from the fact that A and A' are mutually independent and the second inequality comes from the fact that $p_{A,\mathbf{u},\mathbf{u}'} \leq 1$ and $p_{A',\mathbf{u},\mathbf{u}'} \leq 1$. Since $(\alpha_{\hat{A}}, \beta_{\hat{A}})$ satisfies (H1) and (H2), then we have the fact that $(\hat{\mathcal{A}}, p_{\hat{A}})$ has an $(\alpha_{\hat{A}}, \beta_{\hat{A}})$ -hash property. \blacksquare

D. Proof of Lemmas 5 and 6

Let

$$p_{A,\mathbf{u},\mathbf{u}'} \equiv p_A(\{A : A\mathbf{u} = A\mathbf{u}'\})$$

$$p_{B,\mathbf{v},\mathbf{v}'} \equiv p_B(\{A : B\mathbf{v} = A\mathbf{v}'\}).$$

For $(\mathbf{u}', \mathbf{v}') \in \mathcal{T}'$, we have

$$\begin{aligned}
\sum_{\substack{(\mathbf{u},\mathbf{v}) \in \mathcal{T} \\ p_{A,\mathbf{u},\mathbf{u}'} \leq \frac{\alpha_A}{|\text{Im}\mathcal{A}|} \\ p_{B,\mathbf{v},\mathbf{v}'} > \frac{\alpha_B}{|\text{Im}\mathcal{B}|}}} p_{A,\mathbf{u},\mathbf{u}'} p_{B,\mathbf{v},\mathbf{v}'} &= \sum_{\substack{\mathbf{v} \in \mathcal{T}_V \\ p_{B,\mathbf{v},\mathbf{v}'} > \frac{\alpha_B}{|\text{Im}\mathcal{B}|}}} p_{B,\mathbf{v},\mathbf{v}'} \sum_{\substack{\mathbf{u} \in \mathcal{T}_{U|V}(\mathbf{v}) \\ p_{A,\mathbf{u},\mathbf{u}'} \leq \frac{\alpha_A}{|\text{Im}\mathcal{A}|}}} p_{A,\mathbf{u},\mathbf{u}'} \\
&\leq \sum_{\substack{\mathbf{v} \in \mathcal{T}_V \\ p_{B,\mathbf{v},\mathbf{v}'} > \frac{\alpha_B}{|\text{Im}\mathcal{B}|}}} p_{B,\mathbf{v},\mathbf{v}'} \sum_{\substack{\mathbf{u} \in \mathcal{T}_{U|V}(\mathbf{v}) \\ p_{A,\mathbf{u},\mathbf{u}'} \leq \frac{\alpha_A}{|\text{Im}\mathcal{A}|}}} \frac{\alpha_A}{|\text{Im}\mathcal{A}|} \\
&\leq \frac{\left[\max_{\mathbf{v} \in \mathcal{T}_V} |\mathcal{T}_{U|V}(\mathbf{v})| \right] \alpha_A}{|\text{Im}\mathcal{A}|} \sum_{\substack{\mathbf{v} \in \mathcal{T}_V \\ p_{B,\mathbf{v},\mathbf{v}'} > \frac{\alpha_B}{|\text{Im}\mathcal{B}|}}} p_{B,\mathbf{v},\mathbf{v}'} \\
&\leq \frac{\left[\max_{\mathbf{v} \in \mathcal{T}_V} |\mathcal{T}_{U|V}(\mathbf{v})| \right] \alpha_A}{|\text{Im}\mathcal{A}|} \left[\sum_{\substack{\mathbf{v} \in \mathcal{V}^n \setminus \{\mathbf{v}'\} \\ p_{B,\mathbf{v},\mathbf{v}'} > \frac{\alpha_B}{|\text{Im}\mathcal{B}|}}} p_{B,\mathbf{v},\mathbf{v}'} + p_{B,\mathbf{v},\mathbf{v}'} \right] \\
&\leq \frac{\left[\max_{\mathbf{v} \in \mathcal{T}_V} |\mathcal{T}_{U|V}(\mathbf{v})| \right] \alpha_A [\beta_B + 1]}{|\text{Im}\mathcal{A}|}.
\end{aligned} \tag{91}$$

Similarly, we have

$$\sum_{\substack{(\mathbf{u},\mathbf{v}) \in \mathcal{T} \\ p_{A,\mathbf{u},\mathbf{u}'} > \frac{\alpha_A}{|\text{Im}\mathcal{A}|} \\ p_{B,\mathbf{v},\mathbf{v}'} \leq \frac{\alpha_B}{|\text{Im}\mathcal{B}|}}} p_{A,\mathbf{u},\mathbf{u}'} p_{B,\mathbf{v},\mathbf{v}'} \leq \frac{\left[\max_{\mathbf{u} \in \mathcal{T}_U} |\mathcal{T}_{V|U}(\mathbf{u})| \right] \alpha_B [\beta_A + 1]}{|\text{Im}\mathcal{B}|}. \tag{92}$$

for $(\mathbf{u}', \mathbf{v}') \in \mathcal{T}'$. We also have

$$\begin{aligned}
& \sum_{\substack{(\mathbf{u}, \mathbf{v}) \in \mathcal{T} \\ p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} > \frac{\alpha_{\mathbf{A}}}{|\text{Im} \mathcal{A}|} \\ p_{\mathbf{B}, \mathbf{v}, \mathbf{v}'} > \frac{\alpha_{\mathbf{B}}}{|\text{Im} \mathcal{B}|}}} p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} p_{\mathbf{B}, \mathbf{v}, \mathbf{v}'} \\
& \leq |\mathcal{T} \cap \{(\mathbf{u}', \mathbf{v}')\}| + \sum_{\substack{\mathbf{u} \in \mathcal{T}_{\mathcal{U}} \setminus \{\mathbf{u}'\} \\ p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} > \frac{\alpha_{\mathbf{A}}}{|\text{Im} \mathcal{A}|}}} p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} p_{\mathbf{B}, \mathbf{v}, \mathbf{v}'} + \sum_{\substack{\mathbf{v} \in \mathcal{T}_{\mathcal{V}} \setminus \{\mathbf{v}'\} \\ p_{\mathbf{B}, \mathbf{v}, \mathbf{v}'} > \frac{\alpha_{\mathbf{B}}}{|\text{Im} \mathcal{B}|}}} p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} p_{\mathbf{B}, \mathbf{v}, \mathbf{v}'} + \sum_{\substack{(\mathbf{u}, \mathbf{v}) \in \mathcal{T} \\ \mathbf{u} \neq \mathbf{u}', \mathbf{v} \neq \mathbf{v}' \\ p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} > \frac{\alpha_{\mathbf{A}}}{|\text{Im} \mathcal{A}|} \\ p_{\mathbf{B}, \mathbf{v}, \mathbf{v}'} > \frac{\alpha_{\mathbf{B}}}{|\text{Im} \mathcal{B}|}}} p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} p_{\mathbf{B}, \mathbf{v}, \mathbf{v}'} \\
& \leq |\mathcal{T} \cap \{(\mathbf{u}', \mathbf{v}')\}| + \sum_{\substack{\mathbf{u} \in \mathcal{U}^n \setminus \{\mathbf{u}'\} \\ p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} > \frac{\alpha_{\mathbf{A}}}{|\text{Im} \mathcal{A}|}}} p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} + \sum_{\substack{\mathbf{v} \in \mathcal{V}^n \setminus \{\mathbf{v}'\} \\ p_{\mathbf{B}, \mathbf{v}, \mathbf{v}'} > \frac{\alpha_{\mathbf{B}}}{|\text{Im} \mathcal{B}|}}} p_{\mathbf{B}, \mathbf{v}, \mathbf{v}'} + \sum_{\substack{\mathbf{u} \in \mathcal{U}^n \setminus \{\mathbf{u}'\} \\ p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} > \frac{\alpha_{\mathbf{A}}}{|\text{Im} \mathcal{A}|}}} p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} \sum_{\substack{\mathbf{v} \in \mathcal{V}^n \setminus \{\mathbf{v}'\} \\ p_{\mathbf{B}, \mathbf{v}, \mathbf{v}'} > \frac{\alpha_{\mathbf{B}}}{|\text{Im} \mathcal{B}|}}} p_{\mathbf{B}, \mathbf{v}, \mathbf{v}'} \\
& \leq |\mathcal{T} \cap \{(\mathbf{u}', \mathbf{v}')\}| + \beta_{\mathbf{A}} + \beta_{\mathbf{B}} + \beta_{\mathbf{A}} \beta_{\mathbf{B}} \tag{93}
\end{aligned}$$

for $(\mathbf{u}', \mathbf{v}') \in \mathcal{T}'$. Finally, we have

$$\begin{aligned}
& \sum_{\substack{(\mathbf{u}, \mathbf{v}) \in \mathcal{T} \\ (\mathbf{u}', \mathbf{v}') \in \mathcal{T}'}} p_{\mathbf{AB}}(\{(A, B) : A\mathbf{u} = A\mathbf{u}', B\mathbf{v} = B\mathbf{v}'\}) \\
& = \sum_{(\mathbf{u}', \mathbf{v}') \in \mathcal{T}'} \sum_{\substack{(\mathbf{u}, \mathbf{v}) \in \mathcal{T} \\ p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} \leq \frac{\alpha_{\mathbf{A}}}{|\text{Im} \mathcal{A}|} \\ p_{\mathbf{B}, \mathbf{v}, \mathbf{v}'} \leq \frac{\alpha_{\mathbf{B}}}{|\text{Im} \mathcal{B}|}}} p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} p_{\mathbf{B}, \mathbf{v}, \mathbf{v}'} + \sum_{(\mathbf{u}', \mathbf{v}') \in \mathcal{T}'} \sum_{\substack{(\mathbf{u}, \mathbf{v}) \in \mathcal{T} \\ p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} \leq \frac{\alpha_{\mathbf{A}}}{|\text{Im} \mathcal{A}|} \\ p_{\mathbf{B}, \mathbf{v}, \mathbf{v}'} > \frac{\alpha_{\mathbf{B}}}{|\text{Im} \mathcal{B}|}}} p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} p_{\mathbf{B}, \mathbf{v}, \mathbf{v}'} \\
& + \sum_{(\mathbf{u}', \mathbf{v}') \in \mathcal{T}'} \sum_{\substack{(\mathbf{u}, \mathbf{v}) \in \mathcal{T} \\ p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} > \frac{\alpha_{\mathbf{A}}}{|\text{Im} \mathcal{A}|} \\ p_{\mathbf{B}, \mathbf{v}, \mathbf{v}'} \leq \frac{\alpha_{\mathbf{B}}}{|\text{Im} \mathcal{B}|}}} p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} p_{\mathbf{B}, \mathbf{v}, \mathbf{v}'} + \sum_{(\mathbf{u}', \mathbf{v}') \in \mathcal{T}'} \sum_{\substack{(\mathbf{u}, \mathbf{v}) \in \mathcal{T} \\ p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} > \frac{\alpha_{\mathbf{A}}}{|\text{Im} \mathcal{A}|} \\ p_{\mathbf{B}, \mathbf{v}, \mathbf{v}'} > \frac{\alpha_{\mathbf{B}}}{|\text{Im} \mathcal{B}|}}} p_{\mathbf{A}, \mathbf{u}, \mathbf{u}'} p_{\mathbf{B}, \mathbf{v}, \mathbf{v}'} \\
& \leq \sum_{(\mathbf{u}', \mathbf{v}') \in \mathcal{T}'} \frac{|\mathcal{T}| \alpha_{\mathbf{A}} \alpha_{\mathbf{B}}}{|\text{Im} \mathcal{A}| |\text{Im} \mathcal{B}|} + \sum_{(\mathbf{u}', \mathbf{v}') \in \mathcal{T}'} \frac{\left[\max_{\mathbf{v} \in \mathcal{T}_{\mathcal{V}}} |\mathcal{T}_{\mathcal{U}|\mathcal{V}}(\mathbf{v})| \right] \alpha_{\mathbf{A}} [\beta_{\mathbf{B}} + 1]}{|\text{Im} \mathcal{A}|} + \sum_{(\mathbf{u}', \mathbf{v}') \in \mathcal{T}'} \frac{\left[\max_{\mathbf{u} \in \mathcal{T}_{\mathcal{U}}} |\mathcal{T}_{\mathcal{V}|\mathcal{U}}(\mathbf{u})| \right] \alpha_{\mathbf{B}} [\beta_{\mathbf{A}} + 1]}{|\text{Im} \mathcal{B}|} \\
& + \sum_{(\mathbf{u}', \mathbf{v}') \in \mathcal{T}'} [|\mathcal{T} \cap \{(\mathbf{u}', \mathbf{v}')\}| + \beta_{\mathbf{A}} + \beta_{\mathbf{B}} + \beta_{\mathbf{A}} \beta_{\mathbf{B}}] \\
& \leq |\mathcal{T} \cap \mathcal{T}'| + \frac{|\mathcal{T}| |\mathcal{T}'| \alpha_{\mathbf{A}} \alpha_{\mathbf{B}}}{|\text{Im} \mathcal{A}| |\text{Im} \mathcal{B}|} + \frac{|\mathcal{T}'| \left[\max_{\mathbf{v} \in \mathcal{T}_{\mathcal{V}}} |\mathcal{T}_{\mathcal{U}|\mathcal{V}}(\mathbf{v})| \right] \alpha_{\mathbf{A}} [\beta_{\mathbf{B}} + 1]}{|\text{Im} \mathcal{A}|} + \frac{|\mathcal{T}'| \left[\max_{\mathbf{u} \in \mathcal{T}_{\mathcal{U}}} |\mathcal{T}_{\mathcal{V}|\mathcal{U}}(\mathbf{u})| \right] \alpha_{\mathbf{B}} [\beta_{\mathbf{A}} + 1]}{|\text{Im} \mathcal{B}|} \\
& + |\mathcal{T}'| [\beta_{\mathbf{A}} + \beta_{\mathbf{B}} + \beta_{\mathbf{A}} \beta_{\mathbf{B}}], \tag{94}
\end{aligned}$$

where the first inequality comes from (91)–(93). This implies that the joint ensemble $(\mathcal{A} \times \mathcal{B}, p_{\mathbf{AB}})$ satisfies (6) by letting

$$\begin{aligned}
\alpha & \equiv \alpha_{\mathbf{A}} \alpha_{\mathbf{B}} \\
\beta & \equiv \frac{\left[\max_{\mathbf{v} \in \mathcal{T}_{\mathcal{V}}} |\mathcal{T}_{\mathcal{U}|\mathcal{V}}(\mathbf{v})| \right] \alpha_{\mathbf{A}} [\beta_{\mathbf{B}} + 1]}{|\text{Im} \mathcal{A}|} + \frac{\left[\max_{\mathbf{u} \in \mathcal{T}_{\mathcal{U}}} |\mathcal{T}_{\mathcal{V}|\mathcal{U}}(\mathbf{u})| \right] \alpha_{\mathbf{B}} [\beta_{\mathbf{A}} + 1]}{|\text{Im} \mathcal{B}|} + \beta_{\mathbf{A}} + \beta_{\mathbf{B}} + \beta_{\mathbf{A}} \beta_{\mathbf{B}}.
\end{aligned}$$

Then we have Lemmas 5 and 6 from Lemmas 2 and 3, respectively. \blacksquare

E. Proof of Lemma 7

It is easy to show (49) and (50) from the properties (H1) and (H2) of $(\alpha_{A_j}, \beta_{A_j})$ for all $j \in \mathcal{K}$. In the following, we show (47) and (48). Let $p_{\mathbf{u}_j, \mathbf{u}'_j}$ be defined as

$$p_{\mathbf{u}_j, \mathbf{u}'_j} \equiv p_{A_j}(\{A_j : A_j \mathbf{u}_j = A_j \mathbf{u}'_j\}).$$

First, we have

$$\begin{aligned}
& \sum_{\substack{\mathbf{u}_{\mathcal{K}} \in \mathcal{T} \\ p_{\mathbf{u}_j, \mathbf{u}'_j} \leq \frac{\alpha_{A_j}}{|\text{Im} \mathcal{A}_j|} \text{ for all } j \in \mathcal{J} \\ p_{\mathbf{u}_j, \mathbf{u}'_j} > \frac{\alpha_{A_j}}{|\text{Im} \mathcal{A}_j|} \text{ for all } j \in \mathcal{J}^c}} \prod_{j \in \mathcal{K}} p_{\mathbf{u}_j, \mathbf{u}'_j} \\
&= \sum_{\substack{\mathbf{u}_{\mathcal{J}^c} \in \mathcal{T}_{\mathcal{U}_{\mathcal{J}^c}} \\ p_{\mathbf{u}_j, \mathbf{u}'_j} > \frac{\alpha_{A_j}}{|\text{Im} \mathcal{A}_j|}}} \prod_{j \in \mathcal{J}^c} p_{\mathbf{u}_j, \mathbf{u}'_j} \sum_{\substack{\mathbf{u}_{\mathcal{J}} \in \mathcal{T}_{\mathcal{U}_{\mathcal{J}}|\mathcal{U}_{\mathcal{J}^c}}(\mathbf{u}_{\mathcal{J}^c}) \\ p_{\mathbf{u}_j, \mathbf{u}'_j} \leq \frac{\alpha_{A_j}}{|\text{Im} \mathcal{A}_j|}}} \prod_{j \in \mathcal{J}} p_{\mathbf{u}_j, \mathbf{u}'_j} \\
&\leq \left[\max_{\mathbf{u}_{\mathcal{J}^c} \in \mathcal{T}_{\mathcal{U}_{\mathcal{J}^c}}} |\mathcal{T}_{\mathcal{U}_{\mathcal{J}}|\mathcal{U}_{\mathcal{J}^c}}(\mathbf{u}_{\mathcal{J}^c})| \right] \left[\prod_{j \in \mathcal{J}} \frac{\alpha_{A_j}}{|\text{Im} \mathcal{A}_j|} \right] \sum_{\substack{\mathbf{u}_{\mathcal{J}^c} \in \mathcal{T}_{\mathcal{U}_{\mathcal{J}^c}} \\ p_{\mathbf{u}_j, \mathbf{u}'_j} > \frac{\alpha_{A_j}}{|\text{Im} \mathcal{A}_j|}}} \prod_{j \in \mathcal{J}^c} p_{\mathbf{u}_j, \mathbf{u}'_j} \\
&\leq \left[\max_{\mathbf{u}_{\mathcal{J}^c} \in \mathcal{T}_{\mathcal{U}_{\mathcal{J}^c}}} |\mathcal{T}_{\mathcal{U}_{\mathcal{J}}|\mathcal{U}_{\mathcal{J}^c}}(\mathbf{u}_{\mathcal{J}^c})| \right] \left[\prod_{j \in \mathcal{J}} \frac{\alpha_{A_j}}{|\text{Im} \mathcal{A}_j|} \right] \prod_{j \in \mathcal{J}^c} \left[\sum_{\substack{\mathbf{u}_j \in \mathcal{U}_j \setminus \{\mathbf{u}'_j\} \\ p_{\mathbf{u}_j, \mathbf{u}'_j} > \frac{\alpha_{A_j}}{|\text{Im} \mathcal{A}_j|}}} p_{\mathbf{u}_j, \mathbf{u}'_j} + p_{\mathbf{u}'_j, \mathbf{u}'_j} \right] \\
&\leq \left[\max_{\mathbf{u}_{\mathcal{J}^c} \in \mathcal{T}_{\mathcal{U}_{\mathcal{J}^c}}} |\mathcal{T}_{\mathcal{U}_{\mathcal{J}}|\mathcal{U}_{\mathcal{J}^c}}(\mathbf{u}_{\mathcal{J}^c})| \right] \left[\prod_{j \in \mathcal{J}} \frac{\alpha_{A_j}}{|\text{Im} \mathcal{A}_j|} \right] \prod_{j \in \mathcal{J}^c} [\beta_{A_j} + 1] \\
&= \frac{|\mathcal{T}_{\mathcal{J}}|\mathcal{J}^c| \alpha_{A_{\mathcal{J}}} [\beta_{A_{\mathcal{J}^c}} + 1]}{|\text{Im} \mathcal{A}_{\mathcal{J}}|} \tag{95}
\end{aligned}$$

for $\mathbf{u}'_{\mathcal{K}} \in \mathcal{T}'$ and a non-empty set $\mathcal{J} \subsetneq \mathcal{K}$, where the second inequality comes from the property (H3) of (\mathcal{A}_j, p_{A_j}) for all $j \in \mathcal{K}$ and the fact that $p_{\mathbf{u}'_j, \mathbf{u}'_j} = 1$, and the last equality comes from (43) and (44).

Next, we have

$$\begin{aligned}
& \sum_{\substack{\mathbf{u}_{\mathcal{K}} \in \mathcal{T} \setminus \{\mathbf{u}'_{\mathcal{K}}\} \\ p_{\mathbf{u}_j, \mathbf{u}'_j} > \frac{\alpha_{A_j}}{|\text{Im} \mathcal{A}_j|} \text{ for all } j \in \mathcal{K}}} \prod_{j \in \mathcal{K}} p_{\mathbf{u}_j, \mathbf{u}'_j} \leq \sum_{\substack{\mathbf{u}_{\mathcal{K}} \in \mathcal{U}_{\mathcal{K}} \\ p_{\mathbf{u}_j, \mathbf{u}'_j} > \frac{\alpha_{A_j}}{|\text{Im} \mathcal{A}_j|} \text{ for all } j \in \mathcal{K}}} \prod_{j \in \mathcal{K}} p_{\mathbf{u}_j, \mathbf{u}'_j} - \prod_{j \in \mathcal{K}} p_{\mathbf{u}'_j, \mathbf{u}'_j} \\
&= \prod_{j \in \mathcal{J}} \sum_{\substack{\mathbf{u}_j \in \mathcal{U}_j \\ p_{\mathbf{u}_j, \mathbf{u}'_j} > \frac{\alpha_{A_j}}{|\text{Im} \mathcal{A}_j|}}} p_{\mathbf{u}_j, \mathbf{u}'_j} - 1 \\
&= \prod_{j \in \mathcal{J}} \left[\sum_{\substack{\mathbf{u}_j \in \mathcal{U}_j \setminus \{\mathbf{u}'_j\} \\ p_{\mathbf{u}_j, \mathbf{u}'_j} > \frac{\alpha_{A_j}}{|\text{Im} \mathcal{A}_j|}}} p_{\mathbf{u}_j, \mathbf{u}'_j} + p_{\mathbf{u}'_j, \mathbf{u}'_j} \right] - 1 \\
&= \prod_{j \in \mathcal{K}} [\beta_{A_j} + 1] - 1
\end{aligned}$$

$$= \frac{|\mathcal{T}_{\emptyset|\mathcal{K}}| \alpha_{\mathcal{A}_{\emptyset}} [\beta_{\mathcal{A}_{\mathcal{K}}} + 1]}{|\text{Im} \mathcal{A}_{\emptyset}|} - 1 \quad (96)$$

for $\mathbf{u}'_{\mathcal{K}} \in \mathcal{T}'$, where we use the fact that $p_{\mathbf{u}'_j, \mathbf{u}'_j} = 1$ in the second inequality, and the last inequality comes from (43)–(46).

Finally, we have

$$\begin{aligned} & \sum_{\substack{\mathbf{u}_{\mathcal{K}} \in \mathcal{T} \\ \mathbf{u}'_{\mathcal{K}} \in \mathcal{T}'}} p_{\mathcal{A}_{\mathcal{K}}} (\{A_{\mathcal{K}} : A_j \mathbf{u}_j = A_j \mathbf{u}'_j \text{ for all } j \in \mathcal{K}\}) \\ &= \sum_{\mathbf{u}'_{\mathcal{K}} \in \mathcal{T}'} \sum_{\mathbf{u}_{\mathcal{K}} \in \mathcal{T}} \prod_{j=1}^k p_{\mathbf{u}_j, \mathbf{u}'_j} \\ &\leq |\mathcal{T} \cap \mathcal{T}'| + \sum_{\mathbf{u}'_{\mathcal{K}} \in \mathcal{T}'} \sum_{\substack{\mathbf{u}_{\mathcal{K}} \in \mathcal{T} \\ p_{\mathbf{u}_j, \mathbf{u}'_j} \leq \frac{\alpha_{\mathcal{A}_j}}{|\text{Im} \mathcal{A}_j|} \text{ for all } j \in \mathcal{K}}} \prod_{j \in \mathcal{K}} p_{\mathbf{u}_j, \mathbf{u}'_j} + \sum_{\mathbf{u}'_{\mathcal{K}} \in \mathcal{T}'} \sum_{\substack{\mathcal{J} \subsetneq \mathcal{K} \\ \mathcal{J} \neq \emptyset}} \sum_{\substack{\mathbf{u}_{\mathcal{K}} \in \mathcal{T} \\ p_{\mathbf{u}_j, \mathbf{u}'_j} \leq \frac{\alpha_{\mathcal{A}_j}}{|\text{Im} \mathcal{A}_j|} \text{ for all } j \in \mathcal{J} \\ p_{\mathbf{u}_j, \mathbf{u}'_j} > \frac{\alpha_{\mathcal{A}_j}}{|\text{Im} \mathcal{A}_j|} \text{ for all } j \in \mathcal{J}^c}} \prod_{j \in \mathcal{K}} p_{\mathbf{u}_j, \mathbf{u}'_j} \\ &\quad + \sum_{\mathbf{u}'_{\mathcal{K}} \in \mathcal{T}'} \sum_{\substack{\mathbf{u}_{\mathcal{K}} \in \mathcal{T} \setminus \{\mathbf{u}'_{\mathcal{K}}\} \\ p_{\mathbf{u}_j, \mathbf{u}'_j} > \frac{\alpha_{\mathcal{A}_j}}{|\text{Im} \mathcal{A}_j|} \text{ for all } j \in \mathcal{K}}} \prod_{j \in \mathcal{K}} p_{\mathbf{u}_j, \mathbf{u}'_j} \\ &\leq |\mathcal{T} \cap \mathcal{T}'| + \sum_{\mathbf{u}'_{\mathcal{K}} \in \mathcal{T}'} \sum_{\substack{\mathbf{u}_{\mathcal{K}} \in \mathcal{T} \\ p_{\mathbf{u}_j, \mathbf{u}'_j} \leq \frac{\alpha_{\mathcal{A}_j}}{|\text{Im} \mathcal{A}_j|} \text{ for all } j \in \mathcal{K}}} \prod_{j \in \mathcal{K}} \frac{\alpha_{\mathcal{A}_j}}{|\text{Im} \mathcal{A}_j|} + \sum_{\mathbf{u}'_{\mathcal{K}} \in \mathcal{T}'} \sum_{\substack{\mathcal{J} \subsetneq \mathcal{K} \\ \mathcal{J} \neq \emptyset}} \frac{|\mathcal{T}_{\mathcal{J}|\mathcal{J}^c}| \alpha_{\mathcal{A}_{\mathcal{J}}} [\beta_{\mathcal{A}_{\mathcal{J}^c}} + 1]}{|\text{Im} \mathcal{A}_{\mathcal{J}}|} \\ &\quad + \sum_{\mathbf{u}'_{\mathcal{K}} \in \mathcal{T}'} \left[\frac{|\mathcal{T}_{\emptyset|\mathcal{K}}| \alpha_{\mathcal{A}_{\emptyset}} [\beta_{\mathcal{A}_{\mathcal{K}}} + 1]}{|\text{Im} \mathcal{A}_{\emptyset}|} - 1 \right] \\ &\leq |\mathcal{T} \cap \mathcal{T}'| + \frac{|\mathcal{T}| |\mathcal{T}'| \alpha_{\mathcal{A}_{\mathcal{K}}}}{|\text{Im} \mathcal{A}_{\mathcal{K}}|} + |\mathcal{T}'| \left[\sum_{\mathcal{J} \subsetneq \mathcal{K}} \frac{|\mathcal{T}_{\mathcal{J}|\mathcal{J}^c}| \alpha_{\mathcal{A}_{\mathcal{J}}} [\beta_{\mathcal{A}_{\mathcal{J}^c}} + 1]}{|\text{Im} \mathcal{A}_{\mathcal{J}}|} - 1 \right] \quad (97) \end{aligned}$$

for all $\mathcal{T}, \mathcal{T}' \subset \times_{j=1}^k \mathcal{U}_j$, where the first inequality comes from the fact that

$$p_{\mathbf{u}'_j, \mathbf{u}'_j} = 1 > \frac{\alpha_{\mathcal{A}_j}}{|\text{Im} \mathcal{A}_j|}$$

for all $j \in \mathcal{K}$, the second inequality comes from (95) and (96). This implies that the joint ensemble $(\mathcal{A}_{\mathcal{K}}, p_{\mathcal{A}_{\mathcal{K}}})$ satisfies (6) by letting

$$\begin{aligned} \alpha &\equiv \alpha_{\mathcal{A}_{\mathcal{K}}} \\ \beta &\equiv \sum_{\mathcal{J} \subsetneq \mathcal{K}} \frac{|\mathcal{T}_{\mathcal{J}|\mathcal{J}^c}| \alpha_{\mathcal{A}_{\mathcal{J}}} [\beta_{\mathcal{A}_{\mathcal{J}^c}} + 1]}{|\text{Im} \mathcal{A}_{\mathcal{J}}|} - 1 \end{aligned}$$

Then we have (47) and (48) from Lemmas 2 and 3, respectively, where we use the relation

$$\beta_{\mathcal{B}_{\mathcal{K}}} = \frac{|\mathcal{G}_{\emptyset|\mathcal{K}}| \alpha_{\mathcal{A}_{\emptyset}} [\beta_{\mathcal{B}_{\mathcal{K}}} + 1]}{|\text{Im} \mathcal{A}_{\emptyset}|} - 1$$

for the proof of (47). ■

F. Proof of Lemma 8

Since \mathcal{T} is a subset of $\mathcal{T}_{U_{\mathcal{K}}, \gamma}$, we have

$$\mathcal{T}_{\mathcal{U}_{\mathcal{J}^c}} \subset \mathcal{T}_{U_{\mathcal{J}^c}, \gamma}$$

$$\mathcal{T}_{\mathcal{U}_{\mathcal{J}}|\mathcal{U}_{\mathcal{J}^c},\gamma}(\mathbf{u}_{\mathcal{J}^c}) \subset \mathcal{T}_{U_{\mathcal{J}}|U_{\mathcal{J}^c},\gamma}(\mathbf{u}_{\mathcal{J}^c})$$

from Lemma 20. From Lemma 23, we have

$$\begin{aligned} \max_{\mathbf{u}_{\mathcal{J}^c} \in \mathcal{T}_{\mathcal{U}_{\mathcal{J}^c}}} |\mathcal{T}_{\mathcal{U}_{\mathcal{J}}|\mathcal{U}_{\mathcal{J}^c}}(\mathbf{u}_{\mathcal{J}^c})| &\leq \max_{\mathbf{u}_{\mathcal{J}^c} \in \mathcal{T}_{U_{\mathcal{J}^c},\gamma}} |\mathcal{T}_{U_{\mathcal{J}}|U_{\mathcal{J}^c},\gamma}(\mathbf{u}_{\mathcal{J}^c})| \\ &\leq 2^n [H(U_{\mathcal{J}}|U_{\mathcal{J}^c}) + \eta_{\mathcal{U}_{\mathcal{J}}|\mathcal{U}_{\mathcal{J}^c}}(\gamma|\gamma)]. \end{aligned}$$

■

G. Proof of Lemma 9

Before proving the lemma, we prepare some definitions for a convex polytope (see [29]). Let $\mathcal{P} \subset \mathbb{R}^k$ be a convex polytope. Then the linear inequality $\mathbf{c}\mathbf{v} \leq c_0$ is *valid* for \mathcal{P} if this inequality is satisfied for all $\mathbf{v} \in \mathcal{P}$. The *face* of the polytope \mathcal{P} is a set defined by

$$\mathcal{P} \cap \{\mathbf{v} \in \mathbb{R}^k : \mathbf{c}\mathbf{v} = c_0\},$$

where $\mathbf{c}\mathbf{v} \leq c_0$ is valid for \mathcal{P} . If the dimension of a face is 0, it is called a *vertex*.

Lemma 12 ([29, Proposition 2.3(ii)]): Every intersection of the faces of \mathcal{P} is a face of \mathcal{P} .

Lemma 13 ([29, p. 52]): A point $\mathbf{v} \in \mathcal{P}$ cannot be expressed as a convex combination of $\mathcal{P} \setminus \{\mathbf{v}\}$ if and only if it is a vertex of \mathcal{P} .

For $k \geq 1$, let $\mathcal{P}(b^k) \subset \mathbb{R}^{k+1}$ be a convex polytope defined by inequalities

$$v_0 \geq 0 \tag{98}$$

$$v_j \geq 0 \quad \text{for } j \in \mathcal{K} \text{ s.t. } b_j = 0 \tag{99}$$

$$v_j \leq 1 \quad \text{for } j \in \mathcal{K} \text{ s.t. } b_j = 1 \tag{100}$$

$$v_0 + [-1]^{b_j} v_j \leq 1 - b_j \quad \text{for all } j \in \mathcal{K} \tag{101}$$

$$v_0 + \sum_{j \in \mathcal{K}} [-1]^{b_j} v_j \geq 1 - \sum_{j \in \mathcal{K}} b_j \tag{102}$$

where constants $b^k \in \{0, 1\}^k$ and variables v_0, \dots, v_k are considered to be real numbers.

Then we have the following lemmas.

Lemma 14: Let $\mathbf{v} \equiv (v_0, v_1, \dots, v_k)$. Then, for all $j \in \{0\} \cup \mathcal{K}$, inequalities

$$-v_j \leq 0 \tag{103}$$

$$v_j \leq 1 \tag{104}$$

are valid for $\mathcal{P}(b^k)$.

Proof: First, we show (103) and (104) for $j = 0$. The inequality $-v_0 \leq 0$ is valid because it is equivalent to (98). If $b^{(1)} = 0$ then we have

$$v_0 \leq v_0 + v_1 \leq 1$$

from (99) and (101). On the other hand, if $b_j = 1$ then we have

$$v_0 \leq v_1 \leq 1$$

from (100) and (101). Hence we have the fact that $v_0 \leq 1$ is valid for $\mathcal{P}(b^k)$.

Next, we show (103) and (104) for $j \in \mathcal{K}$ satisfying $b_j = 0$. The inequality (103) is valid because it is equivalent to (99). The validity of (104) is obtained by

$$v_j \leq v_0 + v_j \leq 1,$$

which comes from (98) and (101).

Finally, we show (103) and (104) for $j \in \mathcal{K}$ satisfying $b_j = 1$. The inequality (104) is valid because it is equivalent to (100). The validity of (104) is obtained by

$$v_j \geq v_0 \geq 0,$$

which comes from (98) and (101). ■

Lemma 15: For all $b^k \in \{0, 1\}^k$, an integral point \mathbf{v} in $\mathcal{P}(b^k)$ satisfies $\mathbf{v} \in \{0, 1\}^{k+1}$.

Proof: From Lemma 14, the integral point $\mathbf{v} = (v_0, v_1, \dots, v_k)$ satisfies either $v_j = 0$ or $v_j = 1$ for all $j \in \{0\} \cup \mathcal{K}$. ■

Lemma 16: For all $b^k \in \{0, 1\}^k$, $(1, b^k)$ is a member of $\mathcal{P}(b^k)$ and $(0, u^k)$ is a member $\mathcal{P}(b^k)$ for all $u^k \in \{0, 1\}^k$ satisfying $u^k \neq b^k$.

Proof: The first statement $(1, b^k) \in \mathcal{P}(b^k)$ is proved by showing that $\mathbf{v} = (1, b^k)$ satisfies inequalities (98)–(102) for all $b^k \in \{0, 1\}^k$. It is clear that $\mathbf{v} = (1, b^k)$ satisfies (98)–(100). Inequalities (101) and (102) come from the fact that $v_0 = 1$ and b^k satisfies

$$[-1]^{b_j} b_j = -b_j \tag{105}$$

for all $j \in \mathcal{K}$.

The second statement $(0, u^k) \in \mathcal{P}(b^k)$ is proved by showing that $\mathbf{v} = (0, u^k)$ satisfies inequalities (98)–(102) for all $b^k \in \{0, 1\}^k$ and $u^k \in \{0, 1\}^k$ satisfying $u^k \neq b^k$. It is clear that $\mathbf{v} = (0, u^k)$ satisfies (98)–(100) for all $u^k \in \{0, 1\}^k$. Since

$$[-1]^{b_j} u_j = \begin{cases} 1 - b_j, & \text{if } u_j \neq b_j \\ -b_j, & \text{if } u_j = b_j, \end{cases} \tag{106}$$

we have

$$\begin{aligned} 0 + [-1]^{b_j} u_j &\leq \max\{1 - b_j, -b_j\} \\ &\leq 1 - b_j \end{aligned} \tag{107}$$

for all $j \in \mathcal{K}$. This implies that $\mathbf{v} = (0, u^k)$ satisfies (101) for all $j \in \mathcal{K}$. From (106), we have

$$\begin{aligned} \sum_{j \in \mathcal{K}} [-1]^{b_j} u_j &= \sum_{\substack{j \in \mathcal{K} \\ u_j \neq b_j}} [1 - b_j] - \sum_{\substack{j \in \mathcal{K} \\ u_j = b_j}} b_j \\ &= \sum_{\substack{j \in \mathcal{K} \\ u_j \neq b_j}} 1 - \sum_{j \in \mathcal{K}} b_j \\ &\geq 1 - \sum_{j \in \mathcal{K}} b_j, \end{aligned} \tag{108}$$

where the inequality comes from the fact that there is $j' \in \mathcal{K}$ such that $u_{j'} \neq b_{j'}$, because $u^k \neq b^k$. This implies that $\mathbf{v} = (0, u^k)$ satisfies (102). ■

Lemma 17: If $b^k \in \{0, 1\}^k$, then $(0, b^k)$ is not a member of $\mathcal{P}(b^k)$. If $b^k, u^k \in \{0, 1\}^k$ satisfy $u^k \neq b^k$, then $(1, u^k)$ is not a member of $\mathcal{P}(b^k)$.

Proof: First, we show the first statement. From (105), we have

$$\begin{aligned} 0 + \sum_{j \in \mathcal{K}} [-1]^{b_j} b_j &= - \sum_{j \in \mathcal{K}} b_j \\ &< 1 - \sum_{j \in \mathcal{K}} b_j. \end{aligned} \quad (109)$$

This implies that $\mathbf{v} = (0, b^k)$ is not a member of $\mathcal{P}(b^k)$ because it does not satisfy (102).

Next, we show the second statement. Let us assume that $b^k, u^k \in \{0, 1\}^k$ satisfy $u^k \neq b^k$. Then there is $j' \in \mathcal{K}$ such that $u_{j'} \neq b_{j'}$. From (106), we have

$$\begin{aligned} 1 + [-1]^{b_{j'}} u_{j'} &= 1 + 1 - b_{j'} \\ &> 1 - b_{j'}. \end{aligned} \quad (110)$$

Therefore, we have the fact that $\mathbf{v} = (1, u^k)$ is not a member of $\mathcal{P}(b^k)$ because it does not satisfy (101). ■

Lemma 18: The set of all integral vertexes of $\mathcal{P}(b^k)$ is equal to $\mathcal{S}(b^k)$.

Proof: From Lemmas 15–17, we have the fact that all integral points of $\mathcal{P}(b^k)$ are members of $\mathcal{S}(b^k)$ and all members of $\mathcal{S}(b^k)$ are integral points of $\mathcal{P}(b^k)$. In the following, we show that all members of $\mathcal{S}(b^k)$ are vertexes of $\mathcal{P}(b^k)$.

We have

$$\{\mathbf{v}\} = \bigcap_{j=0}^k \{\mathbf{v}' : v'_j = v_j\}$$

for $\mathbf{v} \equiv (v_0, v_1, \dots, v_k) \in \mathcal{S}(b^k) \subset \mathcal{P}(b^k)$. From Lemma 14 and the fact that $\mathbf{v} \in \{0, 1\}^{k+1}$, we have the fact that

$$\begin{aligned} -v'_j &\leq v_j, \text{ for } j \in \mathcal{K} \text{ s.t. } v_j = 0 \\ v'_j &\leq v_j, \text{ for } j \in \mathcal{K} \text{ s.t. } v_j = 1 \end{aligned}$$

are valid inequalities for $\mathcal{P}(b^k)$. This implies that

$$\mathcal{P}(b^k) \cap \{\mathbf{v}' : v'_j = v_j\}$$

is a face of $\mathcal{P}(b^k)$. From Lemma 12, we have the fact that

$$\begin{aligned} \bigcap_{j \in \mathcal{K}} [\mathcal{P}(b^k) \cap \{\mathbf{v}' : v'_j = v_j\}] &= \mathcal{P}(b^k) \cap \bigcap_{j \in \mathcal{K}} \{\mathbf{v}' : v'_j = v_j\} \\ &= \mathcal{P}(b^k) \cap \{\mathbf{v}\} \\ &= \{\mathbf{v}\} \end{aligned} \quad (111)$$

is a face of $\mathcal{P}(b^k)$. Since the dimension of $\{\mathbf{v}\}$ is zero, we have the fact that \mathbf{v} is a vertex of $\mathcal{P}(b^k)$. ■

Lemma 19: For all $\mathbf{v} = (v_0, v_1, \dots, v_k) \in \mathcal{P}(b^k)$ satisfying $0 < v_0 < 1$, \mathbf{v} is not a vertex of $\mathcal{P}(b^k)$.

Proof: Let $u^k \equiv (u^{(1)}, \dots, u^{(k)})$ be defined as

$$u^{(j)} \equiv \frac{v_j - v_0 b^{(j)}}{1 - v_0}.$$

Then \mathbf{v} can be expressed as

$$\mathbf{v} = v_0(1, b^k) + [1 - v_0](0, u^k). \quad (112)$$

Since $\mathbf{v} \neq (0, u^k)$ and $\mathbf{v} \neq (1, b^k)$ for $0 < v_0 < 1$, we have the fact that \mathbf{v} is not a vertex of $\mathcal{P}(b^k)$ from Lemma 13 by assuming that \mathbf{v} , $(1, b^k)$, and $(0, u^k)$ are members of $\mathcal{P}(b^k)$. Then it is enough to show that $(0, u^k) \in \mathcal{P}(b^k)$ by assuming that $\mathbf{v} \in \mathcal{P}(b^k)$ because we have $(1, b^k) \in \mathcal{P}(b^k)$ from Lemma 16.

In the following, v_0 is replaced by 0 and v_j is replaced by $u^{(j)}$ in (98)–(102) when we state that $(0, u^k)$ satisfies inequalities (98)–(102). We show that $(0, u^k)$ satisfies (98)–(102) by assuming $\mathbf{v} \in \mathcal{P}(b^k)$ and $0 < v_0 < 1$.

First, it is clear that $(0, u^k)$ satisfies (98). Next, $(0, u^k)$ satisfies (99) and (100) because

$$u^{(j)} = \frac{v_j - v_0 b^{(j)}}{1 - v_0} = \frac{v_j}{1 - v_0} \geq 0$$

for $j \in \mathcal{K}$ satisfying $b^{(j)} = 0$ and

$$u^{(j)} = \frac{v_j - v_0 b^{(j)}}{1 - v_0} = \frac{v_j - v_0}{1 - v_0} \leq 1,$$

for $j \in \mathcal{K}$ satisfying $b^{(j)} = 1$, where the inequalities come from the fact that $v_0 < 1$ and v_j satisfies (99) and (100). Next, $(0, u^k)$ satisfies (101) because

$$\begin{aligned} 0 + [-1]^{b^{(j)}} u^{(j)} &= \frac{[-1]^{b^{(j)}} [v_j - v_0 b^{(j)}]}{1 - v_0} \\ &\leq \frac{1 - b^{(j)} - v_0 - v_0 [-1]^{b^{(j)}} b^{(j)}}{1 - v_0} \\ &= \frac{[1 - b^{(j)}] [1 - v_0]}{1 - v_0} \\ &= 1 - b^{(j)} \end{aligned} \quad (113)$$

for all $j \in \mathcal{K}$, where the inequality comes from the fact that $v_0 < 1$ and v_j satisfies (101), the second equality comes from (105). Finally, we have the fact that $(0, u^k)$ satisfies (102) because

$$\begin{aligned} 0 + \sum_{j \in \mathcal{K}} [-1]^{b^{(j)}} u^{(j)} &= \sum_{j \in \mathcal{K}} \frac{[-1]^{b^{(j)}} [v_j - v_0 b^{(j)}]}{1 - v_0} \\ &= \sum_{j \in \mathcal{K}} \frac{[-1]^{b^{(j)}} v_j + v_0 b^{(j)}}{1 - v_0} \\ &\geq \frac{1}{1 - v_0} \left[1 - \sum_{j \in \mathcal{K}} b^{(j)} - v_0 + v_0 \sum_{j \in \mathcal{K}} b^{(j)} \right] \\ &= 1 - \sum_{j \in \mathcal{K}} b^{(j)}, \end{aligned} \quad (114)$$

where the second equality comes from (105) and the inequality comes from (102) and the fact that $0 < v_0 < 1$. ■

Now we are in position to prove Lemma 9. From Lemma 18, it is enough to show that there is no non-integral vertex of $\mathcal{P}(b^k)$. Furthermore, from Lemma 19, it is enough to show that there is no non-integral vertex $\mathbf{v} = (v_0, v_1, \dots, v_k)$ of $\mathcal{P}(b^k)$ by assuming $v_0 \in \{0, 1\}$.

Since $\mathcal{P}(b^k) \subset \mathbb{R}^{k+1}$, the vertex of $\mathcal{P}(b^k)$ is determined by $k+1$ equalities representing the face of $\mathcal{P}(b^k)$ described by

$$v_0 = 0 \quad (115)$$

$$v_j = 0 \quad \text{for } j \in \mathcal{K} \text{ s.t. } b_j = 0 \quad (116)$$

$$v_j = 1 \quad \text{for } j \in \mathcal{K} \text{ s.t. } b_j = 1 \quad (117)$$

$$v_0 + [-1]^{b^{(j)}} v_j = 1 - b^{(j)} \quad \text{for } j \in \mathcal{K} \quad (118)$$

$$v_0 + \sum_{j \in \mathcal{K}} [-1]^{b^{(j)}} v_j = 1 - \sum_{j \in \mathcal{K}} b^{(j)}. \quad (119)$$

Since $v_0 \in \{0, 1\}$, we have the fact that $v_j \in \{0, 1\}$ if v_j is determined by one of the equalities (116)–(118).

First, we consider the case where the vertex $\mathbf{v} = (v_0, v_1, \dots, v_k) \in \mathcal{P}(b^k)$ does not satisfy (119). Then \mathbf{v} should be determined by equalities (116)–(118) and we have the fact that $v_j \in \{0, 1\}$ for all $j \in \mathcal{K}$. This implies that \mathbf{v} is an integral point of $\mathcal{P}(b^k)$.

Next, we consider the case where the vertex $\mathbf{v} = (v_0, v_1, \dots, v_k) \in \mathcal{P}(b^k)$ satisfies (119). Since $k-1$ of k variables v_1, \dots, v_k should be determined by equalities (116)–(118), we have the fact that these variables are integers. Then, from (119) and the fact that $v_0 \in \{0, 1\}$, the remaining variable is also an integer. This implies that \mathbf{v} is an integral point of $\mathcal{P}(b^k)$.

Finally, from the above observations, we have the fact that all vertexes $\mathbf{v} = (v_0, v_1, \dots, v_k) \in \mathcal{P}(b^k)$ satisfying $v_0 \in \{0, 1\}$ are integral points. This implies that there is no non-integral vertex $\mathbf{v} = (v_0, v_1, \dots, v_k)$ of $\mathcal{P}(b^k)$ that satisfies $v_0 \in \{0, 1\}$. ■

H. Method of Types

We use the following lemmas for a set of typical sequences.

Lemma 20 ([28, Theorem 2.5][17, Lemma 23]): If $\mathbf{v} \in \mathcal{T}_{V,\gamma}$ and $\mathbf{u} \in \mathcal{T}_{U|V,\gamma'}(\mathbf{v})$, then $(\mathbf{u}, \mathbf{v}) \in \mathcal{T}_{UV,\gamma+\gamma'}$. If $(\mathbf{u}, \mathbf{v}) \in \mathcal{T}_{UV,\gamma}$, then $\mathbf{u} \in \mathcal{T}_{U,\gamma}$ and $\mathbf{u} \in \mathcal{T}_{U|V,\gamma}(\mathbf{v})$.

Lemma 21 ([28, Theorem 2.7][17, Lemma 25]): Let $0 < \gamma \leq 1/8$. Then,

$$\left| \frac{1}{n} \log \frac{1}{\mu_U(\mathbf{u})} - H(U) \right| \leq \zeta_U(\gamma)$$

for all $\mathbf{u} \in \mathcal{T}_{U,\gamma}$, and

$$\left| \frac{1}{n} \log \frac{1}{\mu_{U|V}(\mathbf{u}|\mathbf{v})} - H(U|V) \right| \leq \zeta_{U|V}(\gamma'|\gamma)$$

for $\mathbf{v} \in \mathcal{T}_{V,\gamma}$ and $\mathbf{u} \in \mathcal{T}_{U|V,\gamma'}(\mathbf{v})$, where $\zeta_U(\gamma)$ and $\zeta_{U|V}(\gamma'|\gamma)$ are defined in (2) and (3), respectively.

Lemma 22 ([28, Theorem 2.8][17, Lemma 26]): For any $\gamma > 0$, and $\mathbf{v} \in \mathcal{V}^n$,

$$\begin{aligned} \mu_U([\mathcal{T}_{U,\gamma}]^c) &\leq 2^{-n[\gamma-\lambda_U]} \\ \mu_{U|V}([\mathcal{T}_{U|V,\gamma}(\mathbf{v})]^c|\mathbf{v}) &\leq 2^{-n[\gamma-\lambda_{U|V}]}, \end{aligned}$$

where λ_U and $\lambda_{U|V}$ are defined in (1).

Lemma 23 ([28, Theorem 2.9][17, Lemma 27]): For any $\gamma > 0$, $\gamma' > 0$, and $\mathbf{v} \in \mathcal{T}_{V,\gamma}$,

$$\left| \frac{1}{n} \log |\mathcal{T}_{U,\gamma}| - H(U) \right| \leq \eta_U(\gamma)$$

$$\left| \frac{1}{n} \log |\mathcal{T}_{U|V, \gamma'}(\mathbf{v})| - H(U|V) \right| \leq \eta_{U|V}(\gamma'|\gamma),$$

where $\eta_U(\gamma)$ and $\eta_{U|V}(\gamma'|\gamma)$ are defined in (4) and (5), respectively.

ACKNOWLEDGEMENTS

The authors wish to thank Prof. Uyematsu for introducing the problem of the broadcast channel coding. The authors also wish to thank Prof. Watanabe for introducing the reference [8].

REFERENCES

- [1] A. Bennatan and D. Burshtein, "On the application of LDPC codes to arbitrary discrete-memoryless channels," *IEEE Trans. Inform. Theory*, vol. IT-50, no. 3, pp. 417–438, Mar. 2004.
- [2] P. Berlin and D. Tuninetti, "LDPC codes for fading Gaussian broadcast channels," *IEEE Trans. Inform. Theory*, vol. IT-51, no. 6, pp. 2173–2182, Jun. 2005.
- [3] T. M. Cover, "Broadcast Channels," *IEEE Trans. Inform. Theory*, vol. IT-18, no. 1, pp. 2–14, Jan. 1972.
- [4] T. M. Cover, "A proof of the data compression theorem of Slepian and Wolf for ergodic sources," *IEEE Trans. Inform. Theory*, vol. IT-21, no. 2, pp. 226–228, Mar. 1975.
- [5] I. Csiszár, "Linear codes for sources and source networks: Error exponents, universal coding," *IEEE Trans. Inform. Theory*, vol. IT-28, no. 4, pp. 585–592, Jul. 1982.
- [6] J. L. Carter and M. N. Wegman, "Universal classes of hash functions," *J. Comput. Syst. Sci.*, vol. 18, pp. 143–154, 1979.
- [7] A. El Gamal and E. Van der Meulen, "A proof of Marton's coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inform. Theory*, vol. IT-27, no. 1, pp. 120–122, Jan. 1981.
- [8] A. El Gamal and Y.H. Kim, "Lecture notes on network information theory," available at [arXiv:1001.3404\[cs.IT\]](https://arxiv.org/abs/1001.3404), 2010.
- [9] J. Feldman, M.J. Wainwright, and D.R. Karger, "Using linear programming to decode binary linear codes," *IEEE Trans. Inform. Theory*, vol. IT-51, no. 3, pp. 954–972, Mar. 2005.
- [10] A. A. Gohari and V. Anantharam, "Evaluation of Marton's inner bound for the general broadcast channel," *Proc. 2009 IEEE Int. Symp. Inform. Theory*, Seoul, Korea, Jun. 28–Jul. 3, 2009, pp. 2462–2466.
- [11] B. E. Hajek and M. B. Pursley, "Evaluation of an achievable rate region for the broadcast channel," *IEEE Trans. Inform. Theory*, vol. IT-25, no. 1, pp. 36–46, Jan. 1979.
- [12] A. J. Hoffman and J. B. Kruskal, "Integral boundary points of convex polyhedra," *Linear Inequalities and Related Systems (H.W. Kuhn and A.W. Tucker, Eds.)*, Princeton University Press, pp. 223–246, 1956.
- [13] Y. Liang, G. Kramer, and S. Shamai(Shitz), "Capacity outer bounds for broadcast channels," *Proc. 2008 IEEE Information Theory Workshop*, Porto, Portugal, May 5–9, 2008, pp. 2–4.
- [14] D. J. C. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Trans. Inform. Theory*, vol. IT-45, no. 2, pp. 399–431, Mar. 1999.
- [15] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inform. Theory*, vol. IT-25, no. 3, pp. 306–311, May 1979.
- [16] J. Muramatsu, T. Uyematsu, and T. Wadayama, "Low density parity check matrices for coding of correlated sources," *IEEE Trans. Inform. Theory*, vol. IT-51, no. 10, pp. 3645–3653, Oct. 2005.
- [17] J. Muramatsu and S. Miyake, "Hash property and coding theorems for sparse matrices and maximal-likelihood coding," *IEEE Trans. Inform. Theory*, vol. IT-56, no. 5, pp. 2143–2167, May 2010.
- [18] J. Muramatsu and S. Miyake, "Hash property and fixed-rate universal coding theorems," *IEEE Trans. Inform. Theory*, vol. IT-56, no. 6, pp. 2688–2698, Jun. 2010.
- [19] J. Muramatsu and S. Miyake, "Construction of codes for wiretap channel and secret key agreement from correlated source outputs by using sparse matrices," submitted to *IEEE Trans. Inform. Theory*, available at [arXiv:0903.0414\[cs.IT\]](https://arxiv.org/abs/0903.0414),
- [20] T. Murayama, "Statistical mechanics of data compression theorem," *J. Phys. A: Math. Gen.*, vol. 35: L95L100, 2002.
- [21] C. Nair, "An outer bound for 2-receiver discrete memoryless broadcast channels," available at [arXiv:0807.3593\[cs.IT\]](https://arxiv.org/abs/0807.3593), 2008.
- [22] C. Nair and A. El Gamal, "An outer bound to the capacity region of the broadcast channel," *IEEE Trans. Inform. Theory*, vol. IT-53, no. 1, pp. 350–355, Jan. 2007.
- [23] K. Nakamura, Y. Kabashima, R. Morelos-Zaragoza, and D. Saad, "Statistical mechanics of broadcast channels using low-density parity check codes," *Phys. Rev. E*, vol. 67, no. 036703, 2003.

- [24] T. Ohtsuki, "LDPC codes in communications and broadcasting," *IEICE Trans. Communications.*, vol. E90-B, no. 3, pp. 440–453, Mar. 2007.
- [25] M. Ramezani and M. Ardakani, "Disjoint LDPC coding for Gaussian broadcast channels," *Proc. 2009 IEEE Int. Symp. Inform. Theory*, Seoul, Korea, Jun. 28–Jul. 3, 2009, pp. 938–942.
- [26] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. IT-19, no. 4, pp. 471–480, Jul. 1973.
- [27] D. Schonberg, S. S. Pradhan, and K. Ramchandran, "LDPC codes can approach the Slepian Wolf bound for general binary sources," *40th Annual Allerton Conference on Communication, Control, and Computing*, Allerton House, Monticello, Illinois, Oct. 2002.
- [28] T. Uyematsu, *Gendai Shannon Riron*, Baifukan, 1998 (in Japanese).
- [29] G. M. Ziegler, *Lectures on Polytopes*, Springer Science+Business Media, LLC, 2006.